Generalized TAP Free Energy

WEI-KUO CHEN
University of Minnesota
DMITRY PANCHENKO
University of Toronto
AND
ELIRAN SUBAG
Weizmann Institute of Science

Abstract

We consider the mixed p-spin mean-field spin glass model with Ising spins and investigate its free energy in the spirit of the TAP approach, named after Thouless, Anderson, and Palmer [67]. More precisely, we define and compute the generalized TAP correction, and establish the corresponding generalized TAP representation for the free energy. In connection with physicists' replica theory, we introduce the notion of generalized TAP states, which are the maximizers of the generalized TAP free energy, and show that their order parameters match the order parameter of the ancestor states in the Parisi ansatz. We compute the critical point equations of the TAP free energy that generalize the classical TAP equations for pure states. Furthermore, we give an exact description of the region where the generalized TAP correction is replica symmetric, in which case it coincides with the classical TAP correction, and show that Plefka's condition is necessary for this to happen. In particular, our result shows that the generalized TAP correction is not always replica symmetric on the points corresponding to the Edwards-Anderson parameter. © 2022 Wiley Periodicals, Inc.

1 Introduction

How does a function on a high-dimensional space $\mathbb{R}^N$ (i.e., for large N) typically look? For probabilists, "typically" means endowing some natural space of functions with a probability measure and understanding what occurs with high probability. One natural family of functions consists of random homogeneous polynomials of degree $p \geq 1$ in the coordinates of $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_N) \in \mathbb{R}^N$,

$$H_{N,p}(\boldsymbol{\sigma}) = \frac{1}{N^{(p-1)/2}} \sum_{i_1,\dots,i_p=1}^{N} g_{i_1,\dots,i_p} \sigma_{i_1} \cdots \sigma_{i_p}, \tag{1.1}$$

whose coefficients $g_{i_1,\dots,i_p}$ are independent standard Gaussian variables. In this case, the domain is often restricted to the hypercube $\Sigma_N := \{-1, +1\}^N$ or sphere

Communications on Pure and Applied Mathematics, 0001–0086 (PREPRINT)

$S_N := \{\|\sigma\|_2 = \sqrt{N}\}$, depending on whether the motivation comes from a discrete or continuous setting; the scaling $N^{-(p-1)/2}$ in the definition of $H_{N,p}$ is chosen so that the maximum is typically of order N. More generally, assuming that the processes $H_{N,p}$ above are independent of each other for $p \geq 1$, we will also consider their linear combinations

$$H_N(\sigma) = \sum_{p\geq 1} \beta_p H_{N,p}(\sigma), \tag{1.2}$$

for some sequence $(\beta_p)_{p\geq 1}$ that decreases fast enough, for example, such that $\sum_{p\geq 1} 2^p \beta_p^2 < \infty$.

In statistical physics, the random processes $H_N(\sigma)$ are called spin glass models, or Hamiltonians. A classical example is the Sherrington-Kirkpatrick (SK) model [56], defined by (1.1) with $p = 2$ and $\sigma \in \Sigma_N$. For general $p \geq 1$, $H_{N,p}(\sigma)$ is called the pure p-spin model, and the linear combination $H_N(\sigma)$ is called a mixed p-spin model. In this paper we will work with general mixed models with Ising spins, that is, when the domain of $H_N(\sigma)$ is Σ_N. The spherical case S_N will also be discussed occasionally to put things into a more general context.

Going back to the question we started with, one may ask, for example, what is the maximal value of $H_N(\sigma)$ over Σ_N or S_N, or what is the structure of the set of all near maxima (on the right scale)? More generally, one may wish to know what is the typical cardinality or volume of the set of points where $H_N(\sigma)$ approximately takes a given value, or what is the structure of the same set. However, instead of tackling these questions directly, one often first studies "smooth approximations" of various quantities parametrized by the so-called inverse-temperature parameter $\beta > 0$—an idea common in statistical physics. For example, if for certainty we consider the Ising case Σ_N, relying on (1.6) and (1.7) below, instead of the cardinality of an approximate level set and its geometric structure, one can first study the free energy

$$F_N(\beta) = \frac{1}{N} \log \sum_{\sigma\in\Sigma_N} e^{\beta H_N(\sigma)} \tag{1.3}$$

and the Gibbs measure

$$G_{N,\beta}(A) = \frac{\sum_{\sigma\in A} \exp(\beta H_N(\sigma))}{\sum_{\sigma\in\Sigma_N} \exp(\beta H_N(\sigma))}. \tag{1.4}$$

In the spherical case, the summation is replaced by integration with respect to the Haar measure on S_N. It is well-known that the free energy concentrates around its expectation and that the expectation has a limit $F(\beta) := \lim_{N\to\infty} \mathbb{E} F_N(\beta)$ [29, 45], which is differentiable in $\beta > 0$ [63, 65, 66]. Standard concentration of

measure inequalities then imply that, asymptotically, the maximum of the Hamiltonian (also called the ground state energy, ignoring the minus sign) can be computed via the free energy as

$$E_\star := \lim_{N\to\infty} \max_{\sigma\in\Sigma_N} \frac{H_N(\boldsymbol{\sigma})}{N} = \lim_{\beta\to\infty} \frac{F(\beta)}{\beta}. \tag{1.5}$$

Moreover, it is known [6] that given an energy level $E \in (0, E_\star)$, if we choose β so that $E = F'(\beta)$, the cardinality on the logarithmic scale of the corresponding approximate level set can be expressed via the free energy,

$$\lim_{N\to\infty} \frac{1}{N} \log \#\left\{\boldsymbol{\sigma} : \left|\frac{1}{N} H_N(\boldsymbol{\sigma}) - E\right| < \varepsilon\right\} = F(\beta) - \beta F'(\beta) + O(\varepsilon), \tag{1.6}$$

and the Gibbs measure concentrates on the same set,

$$\lim_{N\to\infty} G_{N,\beta}\left\{\boldsymbol{\sigma} : \left|\frac{1}{N} H_N(\boldsymbol{\sigma}) - E\right| < \varepsilon\right\} = 1. \tag{1.7}$$

In this paper, we will focus on results in the language of the free energies and Gibbs measures for all finite temperatures $\beta < \infty$, and in the follow-up paper [20], we translate these results to the $\beta = \infty$ limit, which concerns near maximizers and their geometry.

To motivate our main results and informally illustrate some of the ideas behind them, consider the following question. Can one identify (in some nontrivial way) points $m \in (-1, 1)^N$ *inside the cube* such that, for small $\varepsilon > 0$, the *narrow band* of configurations $\boldsymbol{\sigma} \in \Sigma_N$ close to the hyperplane perpendicular to m,

$$B(m, \varepsilon) = \left\{\boldsymbol{\sigma} \in \Sigma_N : |R(\boldsymbol{\sigma}, m) - R(m, m)| = \frac{1}{N}|m \cdot (\boldsymbol{\sigma} - m)| < \varepsilon\right\}, \tag{1.8}$$

contains a large number of points with some given energy $\frac{1}{N} H_N(\boldsymbol{\sigma}) \approx E$? As we mentioned above, studying this question means fixing β as in (1.7) above and "large number" means that the Gibbs measure of the set of such points is not too small. It turns out that, without additional structure, such points m are too common to be interesting and, moreover, the "measure" of near maximizers in a band fluctuates too wildly to hope for a meaningful criterion. However, as was pointed out in [59], if we add an additional constraint that there are many *nearly orthogonal* directions $\boldsymbol{\sigma} - m$ inside the band with $\frac{1}{N} H_N(\boldsymbol{\sigma}) \approx E$, then the answer is yes. In fact, for $E \in (0, E_\star)$, such special points m can be characterized through their energy $H_N(m)$ and location m by checking that

$$\frac{\beta}{N} H_N(m) + \mathrm{TAP}_\beta(\mu_m) \approx F(\beta), \tag{1.9}$$

where $\mathrm{TAP}_\beta(\mu_m)$, which we call the *generalized TAP correction* (after Thouless, Anderson, and Palmer), is a *deterministic* function of the empirical measure

$$\mu_m = \frac{1}{N} \sum_{i\le N} \delta_{m_i}. \tag{1.10}$$

Moreover, with high probability over the choice of random coefficients in $H_N(\sigma)$, this criterion can be applied *simultaneously* to all $m \in (-1,1)^N$. As will be explained in the next section, in addition to yielding such a surprisingly simple description of the "special points," the idea of looking at many nearly orthogonal directions has a clear motivation coming from the theory of spin glasses in physics.

While (1.9) will be proved in the current work, in the follow-up work [20] we will deal with the zero temperature $\beta = \infty$ analogue of these results, or, equivalently, the maximal energy value $E = E_\star$. In this case, with appropriate deterministic function TAP_∞, the special points m whose bands have properties analogous to the above are characterized by

$$\frac{1}{N} H_N(m) + \mathrm{TAP}_\infty(\mu_m) \approx E_\star. \tag{1.11}$$

In a recent paper [59], a natural way to define the correction $\mathrm{TAP}_\beta(\mu_m)$ in the spherical-spin case S_N was suggested, from which the following *generalized TAP representation* follows rather quickly: for any $q \in [0,1)$ that belongs to the support of the so-called Parisi measure (see (2.8) below for definition), for large N,

$$F(\beta) \approx \max_{m:\, \|m\|^2 = Nq} \left(\frac{\beta}{N} H_N(m) + \mathrm{TAP}_\beta(\mu_m) \right). \tag{1.12}$$

While the definition in [59] also makes sense for models with Ising spins, it is not obvious at all that this correction can be computed explicitly for those models. In this paper, we introduce some new ideas to solve this problem, and explicitly express the correction via a Parisi-type [51, 52] variational formula. As a result, this yields the generalized TAP representation (1.12) for the Ising case.

Representations of the type (1.9) started from the paper [67], where Thouless, Anderson, and Palmer derived (nonrigorously, using an expansion of the partition function around the local magnetizations) a representation for the free energy of the SK model [56] called the *TAP free energy,*

$$\begin{aligned} F_N(\beta) \approx{}& \frac{\beta}{N} H_N(m) + \frac{\beta^2}{2} \left(1 - \frac{\|m\|_2^2}{N} \right)^2 \\ & - \frac{1}{N} \sum_{i=1}^{N} \left(\frac{1+m_i}{2} \log \frac{1+m_i}{2} + \frac{1-m_i}{2} \log \frac{1-m_i}{2} \right), \end{aligned} \tag{1.13}$$

where $m \in [-1,1]^N$ is some critical point of the right-hand side, $\nabla \mathrm{RHS} = 0$ or

$$m_i = \tanh\left(\left(\beta \nabla H_N(m) \right)_i - 2\beta^2 m_i \left(1 - \frac{\|m\|_2^2}{N} \right) \right) \quad \forall i \le N. \tag{1.14}$$

As the authors of [67] explained, the problem of computing the free energy is then reduced to finding the random solutions m of (1.14), known as the *TAP equations*, subject to a certain convergence condition proposed in [67], and applying (1.13) to m—a problem "not much easier" than the original, in their own words. The representation (1.12) we establish in this paper is a more general analogue of the TAP

free energy representation (1.13), which is well-motivated and fully rigorous. Similarly, by computing the critical point equations for the right-hand side of (1.12), we will derive the *generalized TAP equations* analogous to (1.14). As will be explained in the next section, the motivation comes from the picture that emerged in the subsequent work of physicists in the eighties.

A few years after [67], a real breakthrough was made by Parisi in [51, 52], who discovered the correct formula for the free energy by proposing a very special ansatz within the physicists' replica method. The Parisi solution, which was rather algebraic in nature, was reinterpreted in terms of the geometric structure of the Gibbs measure in the papers by Mézard, Parisi, Sourlas, Toulouse, and Virasoro [35,36], where it was understood, for example, that the ultrametricity of the replica matrix corresponds to ultrametricity of the support of the Gibbs measure in the infinite-volume limit and that the Gibbs measure asymptotically splits into pure states, i.e., disjoint subsets whose structure is simple in an appropriate sense. The so-called order parameter in Parisi's solution, which also plays an important role in the current paper, is a probability measure on $[0, 1]$ called the Parisi measure; for generic models (defined below) it coincides with the asymptotic law of the overlap $R(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) := \frac{1}{N}\sum_{i=1}^{N} \sigma_i^1 \sigma_i^2$ of i.i.d. samples from the Gibbs measure.

The connection between the Parisi ansatz and the classical TAP free energy representation (1.13) was well-understood in the physics literature. In the setting of the SK model, some form of TAP equations for the ancestor states (see below) were derived by Mézard and Virasoro in [38]. However, rigorous mathematical results beyond the high temperature region (see, e.g., [11, 12, 65]) started appearing only more recently. For example, it was confirmed in [19] that the TAP representation of the free energy holds at the level of pure states, that is, a formula of the form of (1.12) holds for q_{EA}, the rightmost point in the support of the Parisi measure. Also, the Mézard-Virasoro equations [38] for mixed p-spin models were derived by Auffinger and Jagannath [7, 8] (see also Remark 2.6 below). Furthermore, recently, Belius and Kistler [9] developed a new method in the setting of the spherical 2-spin model. Lastly, the TAP representation was established at the level of pure states for the spherical pure p-spin models with $p \geq 3$ and $\beta \gg 1$ in [58] and for some spherical mixed p-spin models in the 1-RSB regime with $\beta \gg 1$ by Ben Arous, Zeitouni, and one of the authors in [10]. In the latter works [10, 58], the calculations leading to the TAP representation also yielded certain explicit pure state decompositions, in which each state is centered around a local maximum of the Hamiltonian which also maximizes a certain free energy. More generally, the barycenters, or the so-called local magnetizations, of the abstract pure states decompositions of Talagrand [64] and Jagannath [30] are approximate maximizers m as in (1.12), which correspond to the rightmost point in the support of the Parisi measure.

In this work, we mainly focused on positive temperature $\beta < \infty$ analysis, which deals with energy levels $\frac{1}{N}H_N(\boldsymbol{\sigma}) \approx E$ strictly smaller than the ground state energy $E_\star$. However, the set of near maximizers $\frac{1}{N}H_N(\boldsymbol{\sigma}) \approx E_\star$ also has a rich

and interesting geometry. For example, it was understood, both in the physics literature [37] and rigorously [13, 18, 24], that H_N has exponentially many (in N) near maximizers that are nearly orthogonal to each other. For the spherical models, similar results about the highest critical points are known for the pure models [2, 21, 23, 57, 61] and some mixed models, which are close to being pure [2, 10].

As mentioned above, in [20] we extend our analysis to the zero temperature case $\beta = \infty$. One of the consequences of [20] is that a large set of approximate maximizers of $\frac{1}{N} H_N(m) + \mathrm{TAP}_\infty(\mu_m)$, i.e., approximate generalized TAP solutions, can be arranged in a certain tree structure, whose root is the origin and leaves are points of Σ_N. Since $\mathrm{TAP}_\infty(\mu_{\boldsymbol{\sigma}})$ is constant on Σ_N, the leaves approximately maximize $\frac{1}{N} H_N(\boldsymbol{\sigma})$. This picture is particularly interesting when *full replica symmetry breaking* (FRSB) occurs on the interval $[0, q_{\mathrm{EA}}]$, namely, the support of the Parisi measure is equal to the interval $[0, q_{\mathrm{EA}}]$ as β tends to infinity, which is conjectured to be the case in the SK model; see [37]. In this case, the normalized radii $\|m\|/\sqrt{N}$ of the inner vertices of the tree, which are points $m \in (-1, 1)^N$, are asymptotically dense in $[0, 1]$ and the tree is asymptotically continuous in an appropriate sense. In the spherical case, when the model is FRSB on $[0, q_{\mathrm{EA}}]$, similar insights from [59] inspired an optimization algorithm designed in [60], which outputs a configuration in S_N that roughly maximizes $H_N(\boldsymbol{\sigma})$ in polynomial time in N. In the Ising case, Montanari [39] achieved the same optimization result for the SK model by utilizing the *approximate message passing* (AMP) algorithm based on the TAP equations. Both algorithms start from the origin and iteratively move towards S_N or Σ_N using orthogonal updates, until reaching the approximate optimizer. Montanari proved that his algorithm ends at an approximate TAP solution. In fact, we believe that in each iteration the algorithm jumps from one approximate TAP solution to another. An extension of the algorithm from [39], which optimizes models with FRSB on $[0, q_{\mathrm{EA}}]$ was constructed by El Alaoui, Montanari, and Sellke in [25]. In addition to be useful in the optimizations of the mixed p-spin Hamiltonians, the AMP algorithms driven by the TAP equations have also received great popularity in a number of Baysian inference problems, see, e.g., [34, 40, 41, 68]. Finally, we mention that in another direction, when a certain overlap gap property holds, e.g., for the pure p-spin model with even $p \geq 4$, it was proved that a broad class of algorithms, such as Lipschitzian iteration schemes and low-degree methods fail to produce near ground states for $H_N(\boldsymbol{\sigma})$ in polynomial time in N; see [26,27]. This property is expected to hold generically if the model is not FRSB on $[0, q_{\mathrm{EA}}]$, for instance, when $\beta_2 = 0$ and $\beta_p > 0$ for some $p \geq 3$ (see [3,17,33]).

2 Main Results

2.1 The model

In this paper we will consider the mixed p-spin Hamiltonian $H_N(\boldsymbol{\sigma})$ defined in (1.2) with Ising spins, indexed by $\boldsymbol{\sigma} \in \Sigma_N = \{-1, +1\}^N$. The covariance of the

Gaussian process $H_N(\boldsymbol{\sigma})$ equals

$$\mathbb{E}\, H_N(\boldsymbol{\sigma}^1) H_N(\boldsymbol{\sigma}^2) = N\xi\big(R(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)\big), \tag{2.1}$$

where $R(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2) = \frac{1}{N}\sum_{i=1}^N \sigma_i^1 \sigma_i^2$ is called the overlap of $\boldsymbol{\sigma}^1$ and $\boldsymbol{\sigma}^2$, and where

$$\xi(s) = \sum_{p\geq 1} \beta_p^2 s^p. \tag{2.2}$$

In the Introduction we allowed the (random) free energy and Gibbs measure to depend on an inverse-temperature parameter $\beta > 0$. Of course, β can be absorbed into the coefficients β_p. Hence, to simplify the notation, we redefine the free energy by

$$F_N = \frac{1}{N} \log \sum_{\boldsymbol{\sigma}\in\Sigma_N} \exp H_N(\boldsymbol{\sigma}), \tag{2.3}$$

and the Gibbs measure by

$$G_N(\boldsymbol{\sigma}) = \frac{\exp H_N(\boldsymbol{\sigma})}{\sum_{\boldsymbol{\sigma}\in\Sigma_N} \exp H_N(\boldsymbol{\sigma})}, \tag{2.4}$$

and henceforth use these definitions which do not include β. One can also add an external field term $h\sum_{i\leq N}\sigma_i$ to the Hamiltonian $H_N(\sigma)$, but, for simplicity of notation, we will usually omit it (see also Remark 2.8 below).

A special role will be played by the so-called *generic* mixed p-spin models that satisfy

$$\overline{\mathrm{Span}}\{x^p : \beta_p \neq 0\} = C\big([-1,1], \|\cdot\|_\infty\big), \tag{2.5}$$

which means that sufficiently many of the p-spin terms in the Hamiltonian (1.2) are present in the model.

The limit of the free energy F_N is given by the celebrated Parisi formula [51,52] mentioned above, which was first proved in a seminal work of Talagrand in [62] (building upon a breakthrough by Guerra [28]), and later generalized to models with odd spin interactions in [45]. (The formula for $E_\star$ in (1.5) was derived in [5].) If $\mathcal{M}_{0,1}$ is the space of probability measures on $[0,1]$, for $\zeta \in \mathcal{M}_{0,1}$, let $\Phi_\zeta(t,x)$ be the solution of the Parisi PDE

$$\partial_t \Phi_\zeta = -\frac{\xi''(t)}{2}\big(\partial_{xx}\Phi_\zeta + \zeta(t)(\partial_x \Phi_\zeta)^2\big) \tag{2.6}$$

on $[0,1]\times\mathbb{R}$ with the boundary condition $\Phi_\zeta(1,x) = \log 2\cosh x$. Here $\zeta(s) := \zeta([0,s])$. Define the Parisi functional on $\mathcal{M}_{0,1}$ by

$$\mathcal{P}(\zeta) := \Phi_\zeta(0,0) - \frac{1}{2}\int_0^1 s\xi''(s)\zeta(s)ds. \tag{2.7}$$

Then, the limit of the free energy is given by

$$\lim_{N\to\infty} \mathbb{E} F_N = \inf_{\zeta\in\mathcal{M}_{0,1}} \mathcal{P}(\zeta). \tag{2.8}$$

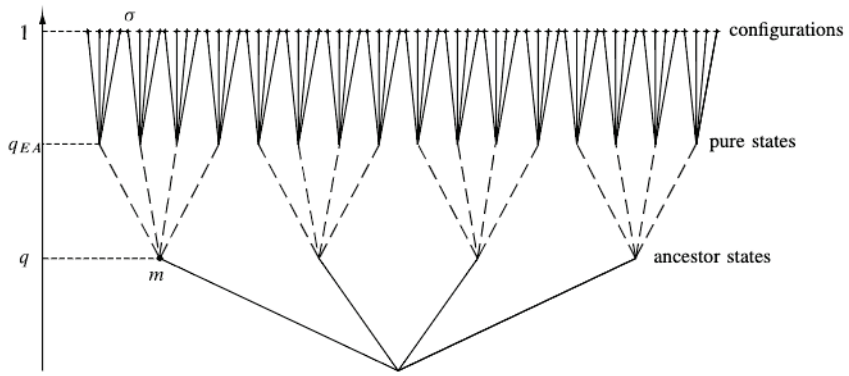

FIGURE 2.1. Ancestor state m on the infinitary tree of states with self-overlap $\frac{1}{N}\|m\|^2 = q \le q_{EA}$.

The minimizer ζ_* is unique (see [4], also [32]) and is called the *Parisi measure*. The solution of the above PDE is usually constructed explicitly for discrete ζ and extended by continuity to all ζ, but one can also show its uniqueness (see [32]).

2.2 Motivation via infinitary nature of the Parisi tree

The Parisi ansatz, schematically depicted in Figure 2.1, states that the Gibbs measure asymptotically decomposes into disjoint pure states, whose magnetizations (barycenters) are organized ultrametrically (see [37]). For simplicity, we plotted only a finite-RSB scenario but, in principle, the overlap can take infinitely many values.

The Parisi ansatz holds for any model that satisfies the Ghirlanda-Guerra identities (see [43]) and, in particular, it holds for generic mixed p-spin models (see section 3.7 in [44]). Since any mixed p-spin model can be approximated by generic models at the level of the free energy, all the results below will apply to nongeneric models as well, and the Parisi ansatz for the generic models will be used as guiding our motivation.

Vertices in the tree in Figure 2.1 below the level of pure states are called *ancestor states*, and they represent branching points when clusters break into smaller subclusters as we zoom in on individual configurations. These ancestor states also have the physical meaning of points m inside the cube

$$m \in [-1, 1]^N. \tag{2.9}$$

(In the spherical models the cube is replaced by the ball of radius $\sqrt{N}$.) The key feature of the Parisi ansatz is that this hierarchical tree of states is *infinitary* in the thermodynamic limit, which means that at each branching point there are infinitely many edges (corresponding to subclusters of a bigger cluster). This infinitary property, in particular, means that all edges in this tree are orthogonal and, for example, a point m corresponding to an ancestor state is perpendicular to $\boldsymbol{\sigma} - m$ for any configuration $\boldsymbol{\sigma}$ (on the scale $1/N$) coming from a pure state that is a descendant

of m. This infinitary property of the Parisi ansatz was used implicitly or explicitly in many applications of ultrametricity (for example, in the proof of ultrametricity itself as well as chaos in temperature in [43, 47], and in the proof of the synchronization mechanism in [46, 50]), and it is central to the main idea in [59] as well as the current paper, which we will explain next.

For m as in (2.9) and $\varepsilon > 0$, we recall the definition of the band centered at m from the introduction,

$$B(m,\varepsilon) = \left\{\boldsymbol{\sigma} \in \Sigma_N : |R(\boldsymbol{\sigma},m) - R(m,m)| = \frac{1}{N}|m\cdot(\boldsymbol{\sigma}-m)| < \varepsilon\right\}. \tag{2.10}$$

If $\varepsilon \geq 2/\sqrt{N}$ then all the bands are nonempty, which can be seen, for example, from Bernstein's inequality: if $\boldsymbol{\sigma}$ comes from the product measure on Σ_N with mean m, then

$$\mathbb{P}\big(|m\cdot(\boldsymbol{\sigma}-m)| \geq N\varepsilon\big) \leq e^{-\frac{N\varepsilon^2}{1+\sqrt{N}\varepsilon/3}} \leq e^{-\frac{4}{1+2/3}} < 1.$$

Given $\delta > 0$ and $n \geq 1$, let us consider a set consisting of n configurations in this narrow band $\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n \in B(m,\varepsilon)$ such that all

$$\widetilde{\boldsymbol{\sigma}}^i = \boldsymbol{\sigma}^i - m \tag{2.11}$$

are almost orthogonal to each other,

$$\begin{aligned} &B_n(m,\varepsilon,\delta) \\ &\quad = \big\{(\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n) \in B(m,\varepsilon)^n : \forall i \neq j,\ \big|R(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^j) - R(m,m)\big| < \delta\big\}. \end{aligned} \tag{2.12}$$

Here, when $n = 1$, this is understood as $B_1(m,\varepsilon,\delta) = B(m,\varepsilon)$. Heuristically, if m corresponds to an ancestor state, then all the descendant pure states are in the band $B(m,\varepsilon)$, so the band carries some nonnegligible weight of the Gibbs measure. Moreover, by the infinitary nature of the tree, we can choose many nearly orthogonal configurations relative to m with nonnegligible Gibbs probability. This means that, for such m, the inequalities

$$F_N \geq \frac{1}{N}\log\sum_{B(m,\varepsilon)} e^{H_N(\boldsymbol{\sigma})} \geq \frac{1}{nN}\log\sum_{B_n(m,\varepsilon,\delta)} e^{\sum_{i=1}^n H_N(\boldsymbol{\sigma}^i)} \tag{2.13}$$

are approximate equalities. Let us introduce the quantity

$$\mathrm{TAP}_{N,n}(m,\varepsilon,\delta) := \frac{1}{nN}\log\sum_{B_n(m,\varepsilon,\delta)} e^{\sum_{i=1}^n [H_N(\boldsymbol{\sigma}^i) - H_N(m)]}, \tag{2.14}$$

which, for simplicity of notation, will often be written with ε and δ omitted,

$$\mathrm{TAP}_{N,n}(m) := \mathrm{TAP}_{N,n}(m,\varepsilon) = \mathrm{TAP}_{N,n}(m,\varepsilon,\delta). \tag{2.15}$$

Then the above inequalities can be rewritten as

$$F_N \geq \frac{H_N(m)}{N} + \mathrm{TAP}_{N,1}(m,\varepsilon) \geq \frac{H_N(m)}{N} + \mathrm{TAP}_{N,n}(m,\varepsilon,\delta). \tag{2.16}$$

Again, for the ancestor states with the self-overlap $\frac{1}{N}\|m\|^2 = q$ corresponding to some q in the support of the Parisi measure, we expect these to be approximate equalities. Moreover, for q in the support of the Parisi measure, one can show that such ancestor states exist, which will imply that

$$F_N \approx \max_{\frac{1}{N}\|m\|^2=q} \left(\frac{H_N(m)}{N} + \mathrm{TAP}_{N,n}(m,\varepsilon,\delta) \right). \tag{2.17}$$

What do we gain by appealing to the infinitary nature of the tree of states in this way?

Given $a, b \in \mathbb{R}$, let $\mathcal{M}_{a,b}$ denote the space of probability measures on $[a,b]$, equipped with the topology of weak convergence. We will always implicitly identify a probability measure μ with its c.d.f., and, for simplicity of notation, write $\mu(x) := \mu((-\infty, x])$. For $\mu, \mu' \in \mathcal{M}_{a,b}$, we will work with the metric

$$d_1(\mu,\mu') = \int_a^b |\mu(x) - \mu'(x)| dx, \tag{2.18}$$

which metrizes weak convergence. We will keep the dependence of d_1 on a and b implicit.

Recall the definition (1.10) of the empirical measure

$$\mu_m = \frac{1}{N} \sum_{i \le N} \delta_{m_i} \in \mathcal{M}_{-1,1}$$

of $m = (m_i)_{i \le N} \in [-1,1]^N$. The key point will be that, for small ε and δ and large n, we can write, with high probability,

$$\mathrm{TAP}_{N,n}(m,\varepsilon,\delta) \approx \mathrm{TAP}(\mu_m) \tag{2.19}$$

uniformly over all m in (2.9), for some specific nonrandom functional TAP: $\mathcal{M}_{-1,1} \to \mathbb{R}$. This functional will be our *generalized TAP correction term* and

$$F_N \approx \max_{\frac{1}{N}\|m\|^2=q} \left(\frac{H_N(m)}{N} + \mathrm{TAP}(\mu_m) \right) \tag{2.20}$$

for q in the support of the Parisi measure is the *generalized TAP representation* of the free energy. Moreover, heuristically, the ancestor states m in the Parisi tree are among the TAP states (near maximizers of the right-hand side), and we will show that these states very much resemble the ancestor states. Let us describe the generalized TAP correction and state our main results precisely.

2.3 Generalized TAP correction and representation

For $\zeta \in \mathcal{M}_{0,1}$, recall the Parisi PDE solution $\Phi_\zeta(q,x)$ from (2.6). Denote the concave conjugate of $\Phi_\zeta(q,\cdot)$ by

$$\Lambda_\zeta(q,a) := \inf_{x \in \mathbb{R}} (\Phi_\zeta(q,x) - ax), \quad a \in [-1,1]. \tag{2.21}$$

It is well-known that $\Phi_\zeta(q,x)$ is a strictly convex function in x and it goes to ± 1 as $x \to \pm\infty$; see [4]. Hence, for each $a \in (-1,1)$, the variational problem defined in $\Lambda_\zeta(q,a)$ has a unique minimizer $\overline{\Psi}(q,a,\zeta)$, which satisfies

$$\partial_x \Phi_\zeta\big(q, \overline{\Psi}(q,a,\zeta)\big) = a, \quad a \in (-1,1). \tag{2.22}$$

With this notation, we can also write

$$\Lambda_\zeta(q,a) = \Phi_\zeta\big(q, \overline{\Psi}(q,a,\zeta)\big) - a\overline{\Psi}(q,a,\zeta), \quad a \in (-1,1). \tag{2.23}$$

We will see that the infimum in (2.21) is finite for $a \in \{-1,1\}$, so the function $\Lambda_\zeta(q,a)$ is continuous on $[-1,1]$. Moreover, since $x \to \Phi_\zeta(q,x)$ is even, so is $a \to \Lambda_\zeta(q,a)$.

For $\mu \in \mathcal{M}_{-1,1}$ such that $\int a^2\, d\mu(a) = q \in [0,1]$, we define

$$\mathrm{TAP}(\mu,\zeta) := \int \Lambda_\zeta(q,a) d\mu(a) - \frac{1}{2}\int_q^1 s\xi''(s)\zeta(s)ds. \tag{2.24}$$

Note that, when $\mu = \delta_1$ (and $q = 1$), the functional is identically equal to zero. Let

$$\mathrm{TAP}(\mu) := \inf_{\zeta \in \mathcal{M}_{0,1}} \mathrm{TAP}(\mu,\zeta). \tag{2.25}$$

Notice that, for a fixed q, this definition depends only on the values of $\zeta(s)$ on the interval $[q,1]$. This means that we could, equivalently, write

$$\mathrm{TAP}(\mu) := \inf_{\zeta \in \mathcal{M}_{q,1}} \mathrm{TAP}(\mu,\zeta), \tag{2.26}$$

where $\mathcal{M}_{q,1}$ is the space of probability distributions on $[q,1]$. We will show in Theorem 3.4 below that the infimum is achieved and the minimizer is unique in $\mathcal{M}_{q,1}$. Because of this, whenever we use the representation (2.25), it will be convenient to use the convention that we minimize over $\zeta \in \mathcal{M}_{0,1}$ such that

$$\zeta(s) = 0 \ \text{ for } s \in [0,q). \tag{2.27}$$

The following is our main result. Recall μ_m in (1.10).

THEOREM 2.1 (Generalized TAP correction). *For any $c,t > 0$, if $\varepsilon,\delta > 0$ are small enough and $n \geq 1$ is large enough, then, for large N,*

$$\mathbb{P}\big(\forall m \in [-1,1]^N : |\mathrm{TAP}_{N,n}(m,\varepsilon,\delta) - \mathrm{TAP}(\mu_m)| < t\big) > 1 - e^{-cN}. \tag{2.28}$$

In particular, we can let $t = t_N$ go to zero slowly with N if we let $\varepsilon = \varepsilon_N$ and $\delta = \delta_N$ go to zero and $n = n_N$ go to infinity slowly enough. Once we computed the TAP correction, we get the TAP representation for the free energy.

THEOREM 2.2 (Generalized TAP representation). *For any q in the support of the Parisi measure of the original model* (1.2)*, in probability,*

$$\lim_{N\to\infty} \left| F_N - \max_{\frac{\|m\|^2}{N}=q} \left(\frac{H_N(m)}{N} + \mathrm{TAP}(\mu_m) \right) \right| = 0. \tag{2.29}$$

Recall that in [19], it was proved that if q_{EA} is the largest point in the support of the Parisi measure for the original Parisi formula of F_N, then

$$\begin{aligned}&\lim_{N\to\infty} F_N\\ &= \lim_{\varepsilon\downarrow 0}\lim_{N\to\infty} \sup_{\substack{m\in[-1,1]^N:\\ \frac{\|m\|_2^2}{N}\in[q_{\mathrm{EA}}-\varepsilon,1]}} \left(\frac{H_N(m)}{N} - \int I(a)d\mu_m(a) + C(q)\right),\end{aligned} \tag{2.30}$$

where

$$\begin{aligned}I(a) &:= \frac{1+a}{2}\log\frac{1+a}{2} + \frac{1-a}{2}\log\frac{1-a}{2},\\ C(q) &:= \frac{1}{2}\big(\xi(1) - \xi(q) - \xi'(q)(1-q)\big).\end{aligned} \tag{2.31}$$

Under an appropriate condition on the empirical measure μ_m, we shall see that the variational formula defined in TAP(μ_m) is solved by the replica symmetric solution in the sense that the minimizer is the Dirac measure at the origin and moreover, our TAP correction term coincides with the sum of the entropy and correction terms in (2.30), that is, $\mathrm{TAP}(\mu_m) = -\int I(a)\mu_m(da) + C(q)$; see Proposition 3.5 below. This allows us to conclude the following version of the classical TAP representation for the free energy from the general representation in Theorem 2.2. We will explain this in more detail and discuss the relation with Plefka's condition [53] in section 3.5.

COROLLARY 2.3 (Classical TAP representation). *The following equation holds almost surely*

$$\lim_{N\to\infty} F_N = \lim_{\varepsilon\downarrow 0}\lim_{N\to\infty} \max\left(\frac{H_N(m)}{N} - \int I(a)d\mu_m(a) + C(q)\right),$$

where the maximum is taken over all vectors $m \in [-1,1]^N$ *that satisfy*

$$\sup_{0\le s\le 1-q} \Gamma_{\mu_m}(s) \le \varepsilon,$$

where $q = \|m\|_2^2/N$ *and* $\Gamma_\mu(s)$ *is defined in* (3.43).

In what follows, we call the near maximizers of the functional

$$\frac{H_N(m)}{N} + \mathrm{TAP}(\mu_m)$$

the *generalized TAP states*. Our definition of the TAP correction was motivated by the fact that ancestor states in the Parisi tree of states should be among the TAP states if $\frac{1}{N}\|m\|^2$ is close to the support of the Parisi measure. Next, we will see that the TAP states have the properties one expects from the ancestor states.

2.4 Properties of generalized TAP states

Let us denote by ζ_m the minimizer in (2.25) or (2.26) (recall our convention (2.27)) corresponding to $m \in [-1,1]^N$ with $q = \frac{1}{N}\|m\|^2$. We will see below that ζ_m has the meaning of the distribution of the overlap for the model on the narrow band $B(m,\varepsilon)$ with its own random external field removed and with a new nonrandom external field added that forces 0 in the support of this distribution ζ_m (see the next section for details). We will show that, if m is a generalized TAP state, then $\zeta_m(s) \approx \zeta_*(s)$ for $s \in [q,1]$, so the order parameters on the band around a TAP state agrees with the Parisi measure of the original model on the interval $[q,1]$. To show this, we will upper-bound the TAP correction by

$$\mathrm{TAP}(\mu) = \inf_{\zeta \in \mathcal{M}_{0,1}} \mathrm{TAP}(\mu,\zeta) \le \mathrm{TAP}(\mu,\zeta_*) \tag{2.32}$$

and obtain the following.

THEOREM 2.4 (TAP states are ancestral). *For any q in the support of the Parisi measure of the original model* (1.2)*, in probability,*

$$\lim_{N\to\infty} \left| F_N - \max_{\frac{\|m\|^2}{N}=q} \left(\frac{H_N(m)}{N} + \mathrm{TAP}(\mu_m,\zeta_*) \right) \right| = 0. \tag{2.33}$$

This together with the representation (2.29) implies that, if m is a TAP state with $q = \frac{1}{N}\|m\|^2 \in \mathrm{supp}(\zeta_*)$ then

$$\mathrm{TAP}(\mu_m,\zeta_m) \approx \mathrm{TAP}(\mu_m,\zeta_*).$$

By continuity properties of the Hamiltonian and the functional $\mathrm{TAP}(\mu,\zeta)$ proved below, this also holds for states with $q = \frac{1}{N}\|m\|^2$ close to the support of ζ_*. We will see (in the proof of Theorem 3.4 below) that the functional $\zeta \to \mathrm{TAP}(\mu,\zeta)$ is d_1-Lipschitz uniformly over μ and has a unique minimizer $\zeta_\mu \in \mathcal{M}_{q,1}$ if $q = \int a^2\, d\mu(a)$, which qualitatively means that

$$\zeta_m \approx \zeta_*; \tag{2.34}$$

i.e., the order parameter ζ_m in the TAP states follows the Parisi measure. This approximation can be quantified, but we do not pursue it here.

Next, in order to describe the critical point equations for the TAP states,

$$\frac{1}{N}\nabla H_N(m) = -\nabla\, \mathrm{TAP}(\mu_m), \tag{2.35}$$

we need to compute the gradient of $\mathrm{TAP}(\mu_m)$. Recall the definition of the function $\overline{\Psi}(q,a,\zeta)$ in (2.22) and let

$$\Psi(q,a,\zeta) := \overline{\Psi}(q,a,\zeta) + a\int_q^1 \xi''(s)\zeta(s)ds. \tag{2.36}$$

The gradient is given by the following formula.

THEOREM 2.5 (Gradient of TAP correction). *For any*

$$m \in (-1,1)^N \quad \text{with } \frac{1}{N}\|m\|^2 = q,$$

if we denote

$$R(m) := \xi''(q)\int_q^1 \zeta_m(s)ds - \int_q^1 \xi''(s)\zeta_m(s)ds, \tag{2.37}$$

then

$$\begin{aligned}\nabla \mathrm{TAP}(\mu_m) &= -\frac{1}{N}\big(\Psi(q, m_i, \zeta_m) + R(m)m_i\big)_{i\le N} \\ &= -\frac{1}{N}\left(\overline{\Psi}(q, m_i, \zeta_m) + m_i\xi''(q)\int_q^1 \zeta_m(s)ds\right)_{i\le N}.\end{aligned} \tag{2.38}$$

Remark 2.6 (Generalized TAP equations). Let us show how (2.35) and (2.38) lead to the generalized TAP equations. If we combine (2.35) and (2.38), we can write

$$\big(\nabla H_N(m)\big)_i - m_i\xi''(q)\int_q^1 \zeta_m(s)ds = \overline{\Psi}(q, m_i, \zeta_m).$$

If we plug both sides into $\partial_x \Phi_{\zeta_m}(q,\cdot)$ and recall the definition of $\overline{\Psi}$, we get

$$\partial_x \Phi_{\zeta_m}\left(q, (\nabla H_N(m))_i - m_i\xi''(q)\int_q^1 \zeta_m(s)ds\right) = m_i. \tag{2.39}$$

These are the TAP equations for generalized TAP states. To compare them with classical equations, we can use that TAP states with $\frac{1}{N}\|m\|^2 = q \in \mathrm{supp}(\zeta_*)$ (or close to the support) must have the order parameter $\zeta_m \approx \zeta_*$, which yields the approximate TAP equations,

$$\partial_x \Phi_{\zeta_*}\left(q, (\nabla H_N(m))_i - m_i\xi''(q)\int_q^1 \zeta_*(s)ds\right) \approx m_i. \tag{2.40}$$

We will discuss the replica symmetric case of TAP correction in the next section in much more detail, but notice that, when $\zeta_* = 1$ for $s \in [q,1]$, (2.40) reduces to

$$\tanh\big((\nabla H_N(m))_i - m_i\xi''(q)(1-q)\big) \approx m_i. \tag{2.41}$$

For the SK model, these are the classical TAP equations, which also appeared in the physics literature for the pure p-spin model in [54] (see also [22]). The equations (2.39), (2.40) are, thus, an extension of the classical TAP equations to all generalized TAP states, for all mixed p-spin models.

Let us remark that (2.39), (2.40) above are self-consistent TAP equations in the sense that they relate state magnetization to itself. They are different from the Mézard-Virasoro equations for the ancestor states derived rigorously in theorem 1.4 in [7], which relate magnetization to the local field, although for the 2-spin SK model considered in Mézard-Virasoro [38], they happen to coincide with (2.40). The main reason is because, in those results, in place of the term $\nabla H_N(m)$

in (2.39), one has the cavity field process at an ancestor state m, which represents the average of $\nabla H_N(\boldsymbol{\sigma})$ over many mutually orthogonal directions around m. For example, if the entire system is in a pure state, the cavity field process corresponds to $\langle \nabla H_N(\boldsymbol{\sigma}) \rangle$, which coincides with $\nabla H_N(\langle \sigma \rangle) = \nabla H_N(m)$ only for the pure 2-spin model, when the gradient $\nabla H_N(\boldsymbol{\sigma})$ does not include interaction terms. This is why in those results one has the term $m_i \int_q^1 \xi''(s)\zeta_*(s)ds$ instead of $m_i\xi''(q)\int_q^1 \zeta_*(s)ds$ in (2.40), and correspondingly $m_i(\xi'(1)-\xi'(q))$ instead of $m_i\xi''(q)(1-q)$ in (2.41), which appears in self-consistent TAP equations, as, e.g., in [54].

Remark 2.7 (Spherical gradient). The formula (2.38) implies that in the spherical directions,

$$\nabla \mathrm{TAP}(\mu_m) \cdot v = -\frac{1}{N}\big(\Psi(q, m_i, \zeta_m)\big)_{i\le N} \cdot v \quad \text{for all } v \perp m. \tag{2.42}$$

There is a physical argument for the formula (2.42) to be satisfied by the ancestor states m in the Parisi ansatz. We will see below that $(\Psi(q, m_i, \zeta_m))$ has the meaning of the unique external field that forces the model on the narrow band $B(m, \varepsilon)$ (with its own random external field removed) to have many orthogonal pure states relative to m. On the other hand, $\nabla H_N(m)$ is the external field of the original model restricted to the band and, when m is an ancestor state, we know that there exist many orthogonal states on the band. This suggests that $\nabla H_N(m)\cdot v = (\Psi(q, m_i, \zeta_m))\cdot v$ for such m and $v \perp m$, which agrees with (2.35) and (2.42). □

Remark 2.8 (Model with external field). One can include an external field to the original model and consider the model with the Hamiltonian

$$H_N^{\mathrm{ef}}(\boldsymbol{\sigma}) = H_N(\boldsymbol{\sigma}) + h\sum_{i=1}^N \sigma_i. \tag{2.43}$$

It will be clear from the discussion below that the TAP correction $\mathrm{TAP}(\mu_m)$ is the same whether or not the external field is present. In fact, this will be, in some sense, a big part of the motivation for our definition of the generalized TAP correction. The only difference will be at the level of the TAP representation (2.29), which will become

$$F_N \approx \max_{\frac{\|m\|^2}{N}=q}\left(\frac{H_N(m)}{N} + h\sum_{i=1}^N m_i + \mathrm{TAP}(\mu_m)\right), \tag{2.44}$$

for q in the support of the Parisi measure of the model with external field h. For simplicity of notation, we will work without the external field, because only trivial modifications are necessary in the case with external field. □

In the next section, we will give an outline of the main ideas in the proof and state further results. For example, we will show that the replica symmetric case of the above generalized TAP correction (when the minimizer in (2.25) equals δ_0) reduces

to the classical TAP correction and give a precise characterization for when that happens. We will show how this implies the necessity of Plefka's condition and, in particular, that the generalized TAP correction is not always replica symmetric on points corresponding to the Edwards-Anderson parameter (see Remark 3.7).

3 General Outline and Further Results

3.1 Utilizing many orthogonal directions

Let us explain the main ideas from [59] that allow us to make the uniform claim (2.28) and at the same time compute things explicitly despite the dependence on large n. This will also allow us to introduce some necessary definitions and notation. We will see that for large n and small ε and δ, the following properties hold:

(a) with high probability, $\mathrm{TAP}_{N,n}(m)$ in (2.14) is close to its expectation, uniformly over all m in (2.9);
(b) adding or removing an external field term in $\mathbb{E}\,\mathrm{TAP}_{N,n}(m)$ has a negligible effect.

Let us explain what these properties mean and sketch why they hold. First of all, let us compute the covariance of the process $H_N(\boldsymbol{\sigma}) - H_N(m)$ for $\boldsymbol{\sigma}$ in the narrow band $B(m,\varepsilon)$,

$$\begin{aligned}\frac{1}{N}\mathbb{E}\big(H_N(\boldsymbol{\sigma}^i) - H_N(m)\big)\big(H_N(\boldsymbol{\sigma}^j) - H_N(m)\big)\\ = \xi(R(\boldsymbol{\sigma}^i,\boldsymbol{\sigma}^j)) - \xi(R(\boldsymbol{\sigma}^i,m)) - \xi(R(\boldsymbol{\sigma}^j,m)) + \xi(R(m,m)).\end{aligned} \tag{3.1}$$

For $(\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n) \in B_n(m,\varepsilon,\delta)$, the covariance for $i \neq j$ is small, by the definition of $B(m,\varepsilon)$ and $B_n(m,\varepsilon,\delta)$. Therefore, the variance of $\sum_{i=1}^n [H_N(\boldsymbol{\sigma}^i) - H_N(m)]$ in (2.14) is roughly of the order nN and, by the Gaussian concentration, the fluctuations of $\mathrm{TAP}_{N,n}(m)$ are of order $(nN)^{-1/2}$ (Lemma 4.3 contains a precise statement). The extra factor $n^{-1/2}$ with large n will allow us to discretize and apply a union bound uniformly over m, implying the first property (a).

To explain the second property (b), it is convenient to think of $H_N(\boldsymbol{\sigma}) - H_N(m)$ as a new mixed p-spin model on the narrow band, as follows. If $q = R(m,m) = \|m\|^2/N$, then, for $\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2 \in B(m,\varepsilon)$,

$$|R(\widetilde{\boldsymbol{\sigma}}^j, m)| = |R(\boldsymbol{\sigma}^j, m) - q| < \varepsilon \tag{3.2}$$

and

$$|R(\boldsymbol{\sigma}^1,\boldsymbol{\sigma}^2) - (R(\widetilde{\boldsymbol{\sigma}}^1,\widetilde{\boldsymbol{\sigma}}^2) + q)| = |R(\widetilde{\boldsymbol{\sigma}}^1, m) + R(\widetilde{\boldsymbol{\sigma}}^2, m)| < 2\varepsilon. \tag{3.3}$$

Therefore, up to the error of order $O(\varepsilon)$, the covariance in (3.1) is approximated by

$$\xi(R(\widetilde{\boldsymbol{\sigma}}^1,\widetilde{\boldsymbol{\sigma}}^2) + q) - \xi(q) = \widehat{\xi}_q(R(\widetilde{\boldsymbol{\sigma}}^1,\widetilde{\boldsymbol{\sigma}}^2)), \tag{3.4}$$

where

$$\hat{\xi}_q(s) := \xi(s+q) - \xi(q) = \sum_{k\geq 1} \beta_k(q)^2 s^k \tag{3.5}$$

and where

$$\beta_k(q)^2 = \sum_{p\geq k} \binom{p}{k} \beta_p^2 q^{p-k}. \tag{3.6}$$

As a result, if $\hat{H}_N^m(\tilde{\boldsymbol{\sigma}})$ is the mixed p-spin Hamiltonian indexed by $\tilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - m$ for $\boldsymbol{\sigma} \in B(m,\varepsilon)$ with the covariance given by

$$\mathbb{E}\hat{H}_N^m(\tilde{\boldsymbol{\sigma}}^1)\hat{H}_N^m(\tilde{\boldsymbol{\sigma}}^2) = N\hat{\xi}_q\big(R(\tilde{\boldsymbol{\sigma}}^1, \tilde{\boldsymbol{\sigma}}^2)\big), \tag{3.7}$$

one can show that (see Lemma 6.2 below)

$$\mathbb{E}\,\mathrm{TAP}_{N,n}(m,\varepsilon,\delta) = \frac{1}{nN}\mathbb{E}\log \sum_{B_n(m,\varepsilon,\delta)} e^{\sum_{\ell=1}^n \hat{H}_N^m(\tilde{\boldsymbol{\sigma}}^\ell)} + O\left(\frac{\varepsilon}{\sqrt{q}}\right) \tag{3.8}$$

uniformly over n, if $q = R(m,m) = \|m\|^2/N > 0$. The case when q is small and m is close to zero will be handled slightly differently, by working with the original model with the external field removed, without any recentering.

The external field term that we mentioned in the property (b) is present in the model $\hat{H}_N^m(\tilde{\boldsymbol{\sigma}})$, because $\beta_1(q) \neq 0$ in (3.5). Define the function ξ_q similarly to (3.5), only with the summation starting from $k = 2$,

$$\xi_q(s) := \xi(s+q) - \xi(q) - \xi'(q)s = \sum_{k\geq 2} \beta_k(q)^2 s^k, \tag{3.9}$$

and let $H_N^m(\tilde{\boldsymbol{\sigma}})$ be the Hamiltonian with the covariance

$$\mathbb{E}H_N^m(\tilde{\boldsymbol{\sigma}}^1)H_N^m(\tilde{\boldsymbol{\sigma}}^2) = N\xi_q\big(R(\tilde{\boldsymbol{\sigma}}^1, \tilde{\boldsymbol{\sigma}}^2)\big), \tag{3.10}$$

with $q = \|m\|^2/N$. Then, in distribution,

$$\hat{H}_N^m(\tilde{\boldsymbol{\sigma}}) = H_N^m(\tilde{\boldsymbol{\sigma}}) + \beta_1(q)\tilde{\boldsymbol{\sigma}} \cdot g,$$

where g is a standard Gaussian vector. Hence, using the pairwise near-orthogonality of $(\tilde{\boldsymbol{\sigma}}^1, \ldots, \tilde{\boldsymbol{\sigma}}^n)$ for $(\boldsymbol{\sigma}^1, \ldots, \boldsymbol{\sigma}^n) \in B_n(m,\varepsilon,\delta)$, namely,

$$\sup_{B_n(m,\varepsilon,\delta)} \left\| \frac{1}{n\sqrt{N}} \sum_{\ell=1}^n \tilde{\boldsymbol{\sigma}}^\ell \right\| \leq \left(\frac{1}{n} + \delta\right)^{1/2},$$

we get that

$$\frac{1}{nN}\left|\mathbb{E}\log \sum_{B_n(m,\varepsilon,\delta)} e^{\sum_{\ell=1}^n \hat{H}_N^m(\tilde{\boldsymbol{\sigma}}^\ell)} - \mathbb{E}\log \sum_{B_n(m,\varepsilon,\delta)} e^{\sum_{\ell=1}^n H_N^m(\tilde{\boldsymbol{\sigma}}^\ell)}\right| \leq \beta_1(q)\left(\frac{1}{n} + \delta\right)^{1/2}.$$

Together with (3.8), this shows that

$$\mathbb{E}\,\mathrm{TAP}_{N,n}(m,\varepsilon,\delta) = \mathbb{E}\,F_{N,n}(m,\varepsilon,\delta) + O\Big(\frac{\varepsilon}{\sqrt{q}} + \Big(\frac{1}{n}+\delta\Big)^{1/2}\Big), \tag{3.11}$$

where we introduce the notation

$$F_{N,n}(m,\varepsilon,\delta) := \frac{1}{nN}\log \sum_{B_n(m,\varepsilon,\delta)} e^{\sum_{\ell=1}^n H_N^m(\widetilde{\boldsymbol{\sigma}}^\ell)} \tag{3.12}$$

for the free energy in the replicated band, with the external field removed.

By the same argument, we may also add a deterministic external field term $h = (h_i)_{i\le N}$, as long as we keep $\|h\|/\sqrt{N}$ bounded. Namely, if we define

$$F_{N,n}^h(m,\varepsilon,\delta) := \frac{1}{nN}\log \sum_{B_n(m,\varepsilon,\delta)} e^{\sum_{\ell=1}^n \left[H_N^m(\widetilde{\boldsymbol{\sigma}}^\ell)+\sum_{i=1}^N h_i\widetilde{\sigma}_i^\ell\right]}, \tag{3.13}$$

then

$$\big|F_{N,n}(m,\varepsilon,\delta) - F_{N,n}^h(m,\varepsilon,\delta)\big| \le \frac{\|h\|}{\sqrt{N}}\Big(\frac{1}{n}+\delta\Big)^{1/2}. \tag{3.14}$$

In other words,

$$\mathbb{E}\,\mathrm{TAP}_{N,n}(m,\varepsilon,\delta) \approx \mathbb{E}\,F_{N,n}^h(m,\varepsilon,\delta),$$

when ε,δ are small, n is large, and $\|h\|/\sqrt{N}$ stays bounded. As with $\mathrm{TAP}_{N,n}(m)$, for simplicity of notation, we will often omit ε and δ and write

$$F_{N,n}^h(m) := F_{N,n}^h(m,\varepsilon) = F_{N,n}^h(m,\varepsilon,\delta). \tag{3.15}$$

Let us sketch how the properties (a) and (b) lead to an explicit calculation of the generalized correction term $\mathrm{TAP}(\mu_m)$, and also contrast what happens in the spherical models vs. models with Ising spins.

3.2 Spherical vs. Ising spin models

Let us start with an overview of the spherical model that was considered in [59], where of course Σ_N should be replaced by S_N, and the sums over configurations should be replaced by integrals. Using the above heuristics of introducing many orthogonal constraints (or the infinitary nature of the tree of states), we get that, uniformly over m,

$$\mathrm{TAP}_{N,n}(m) \approx \mathbb{E}\,\mathrm{TAP}_{N,n}(m) \approx \mathbb{E}\,F_{N,n}(m). \tag{3.16}$$

Since a narrow band on the sphere looks the same for all m with $\frac{1}{N}\|m\|^2 = q$, the right-hand side depends on m only through q. We can see that the constraints $(\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n) \in B(m,\varepsilon,\delta)$ in

$$\mathbb{E}\,F_{N,n}(m) = \frac{1}{nN}\mathbb{E}\log \int_{B_n(m,\varepsilon,\delta)} e^{\sum_{\ell=1}^n H_N^m(\widetilde{\boldsymbol{\sigma}}^\ell)} d\boldsymbol{\sigma}^1\cdots d\boldsymbol{\sigma}^n \tag{3.17}$$

can be expressed by saying that

$$\frac{1}{N}\|\widetilde{\boldsymbol{\sigma}}^i\|^2 = \frac{1}{N}\|\boldsymbol{\sigma}^i - m\|^2 \approx 1 - q$$

and all $\widetilde{\boldsymbol{\sigma}}^1, \ldots, \widetilde{\boldsymbol{\sigma}}^n$ are almost orthogonal to each other, $R(\widetilde{\boldsymbol{\sigma}}^i, \widetilde{\boldsymbol{\sigma}}^j) \approx 0$ for $i \neq j$. Moreover, the narrow band is, essentially, a sphere in one dimension less, so we can think of $H_N^m(\widetilde{\boldsymbol{\sigma}})$ as a new spherical model. The fact that the external field has been removed implies a well-known fact that zero is in the support of the Parisi measure of this new spherical model and, in particular, the overlap constraints $R(\widetilde{\boldsymbol{\sigma}}^i, \widetilde{\boldsymbol{\sigma}}^j) \approx 0$ cannot have a free energy cost. In other words, one can show that

$$\mathbb{E}F_{N,n}(m) \approx \mathbb{E}F_{N,1}(m), \tag{3.18}$$

which is the free energy of a spherical model that can be written as a spherical analogue of the Parisi formula, the Crisanti-Sommers formula. Thus, the generalized TAP correction has a particularly simple form in the spherical models, and this has important consequences, as was demonstrated in [59].

In the Ising spin models, the situation is quite different. First of all, the narrow band depends on m in a complicated way, and the constraint $\boldsymbol{\sigma} \in B(m, \varepsilon)$ can be viewed as a constraint on $\widetilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - m$ of the form

$$\begin{aligned}\widetilde{\boldsymbol{\sigma}} \in \widetilde{B}(m, \varepsilon) &:= B(m, \varepsilon) - m \\ &\approx (\Sigma_N - m) \bigcap \left\{ \frac{1}{N}\|\widetilde{\boldsymbol{\sigma}}\|^2 \approx 1 - q \right\}.\end{aligned} \tag{3.19}$$

So, the first question is: for $n = 1$, can we compute the analogue of the Parisi formula for the free energy $\mathbb{E}F_{N,1}(m)$ on the narrow band? The answer is yes, but this will require some work. However, the bigger issue is that, even if we can compute this free energy, removing the external field term will not result in zero being in the support of the Parisi measure, because of the inherent asymmetry of the band (3.19), and will not allow us to make the step (3.18). The solution to this will be to *add a new external field* to balance out the asymmetry of the band. In other words, using (3.14), we will introduce an external field $h = (h_i)$ at the step (3.16),

$$\mathbb{E}\,\mathrm{TAP}_{N,n}(m) \approx \mathbb{E}F_{N,n}(m) \approx \mathbb{E}F_{N,n}^h(m), \tag{3.20}$$

and, with the right choice of h, we will show that zero is in the support of the Parisi measure of the model on the band and, therefore,

$$\mathbb{E}F_{N,n}^h(m) \approx \mathbb{E}F_{N,1}^h(m). \tag{3.21}$$

The analogue of the Parisi formula for the right-hand side will be our TAP correction $\mathbf{T}(\mu_m)$. The ideas behind finding the right choice of h will be explained below.

Remark 3.1. Notice that the functional $a \to \Lambda_\zeta(q,a)$ in (2.21) is even, and therefore the functional in (2.24) has the symmetry

$$\mathrm{TAP}(\mu_m,\zeta) = \mathrm{TAP}(\mu_{|m|},\zeta),$$

where $\mu_{|m|} = \frac{1}{N}\sum_{i\le N}\delta_{|m_i|}$. Under the transformation $\sigma_i \to \mathrm{sgn}(m_i)\sigma_i$, the overlap between two configurations does not change, while the covariance of all the Hamiltonians as well as definition of the bands depend only on the overlaps. Furthermore, throughout the paper, we will always work with external fields of the form $h_i = \mathrm{sgn}(m_i)v(|m_i|)$ that depend on the coordinates m_i in an antisymmetric fashion, which means that the external field $h_i\widetilde{\sigma}_i$ will also be invariant under this transformation. Because of this, from now on, we can and will assume that

$$m \in [0,1]^N, \tag{3.22}$$

that is, all the coordinates $m_i \ge 0$. In particular, we assume that

$$\mu_m = \frac{1}{N}\sum_{i\le N}\delta_{m_i} \in \mathcal{M}_{0,1}. \tag{3.23}$$

For the rest of the paper, we will work with this definition of $\mathcal{M}_{0,1}$.

3.3 Parisi formula on the band

Let us now state the analogue of the Parisi formula on the band $B(m,\varepsilon)$ for $m \in [0,1]^N$ with the general external fields of the form $h_i = v(m_i)$ for $v \in C([0,1])$, which will be sufficient for our purposes. Since the self-overlap of the configurations $\widetilde{\sigma} = \sigma - m$ on the narrow band is close to $1-q$, it will be natural to work with the space $\mathcal{M}_{0,1-q}$ of all distributions on $[0,1-q]$. Recall the function ξ_q in (3.9) and, for each $a \in [0,1]$ and $\zeta \in \mathcal{M}_{0,1-q}$, denote by $\Phi_{a,\zeta}(t,x)$ the solution of the Parisi PDE on $[0,1-q]\times\mathbb{R}$,

$$\partial_t\Phi_{a,\zeta} = -\frac{\xi_q''(t)}{2}\Big(\partial_{xx}\Phi_{a,\zeta} + \zeta(t)\big(\partial_x\Phi_{a,\zeta}\big)^2\Big) \tag{3.24}$$

with the boundary condition

$$\Phi_{a,\zeta}(1-q,x) = \log 2 - ax + \log\cosh x = \log\sum_{\sigma=\pm1} e^{(\sigma-a)x}. \tag{3.25}$$

Let us define the function $\Psi(a,\zeta)$ by

$$\partial_x\Phi_{a,\zeta}\big(0,\Psi(a,\zeta)\big) = 0, \quad a \in [0,1). \tag{3.26}$$

Note that $\Psi(a,\zeta)$ is well-defined as $\partial_x\Phi_{a,\zeta}(0,\cdot)$ is strictly increasing with

$$\partial_x\Phi_{a,\zeta}(0,\pm\infty) = \pm1 - a.$$

Also, note that $\Phi_{a,\zeta}$ depends on q, but we will keep this dependence implicit for simplicity of notation.

For $\mu \in \mathcal{M}_{0,1}$, $v \in C([0,1])$, and $(\lambda, \zeta) \in \mathbb{R} \times \mathcal{M}_{0,1-q}$, let

$$\mathcal{P}^v_\mu(\lambda,\zeta) = \int \Phi_{a,\zeta}(0, \lambda a + v(a)) d\mu(a) - \frac{1}{2}\int_0^{1-q} s\xi_q''(s)\zeta(s)ds. \tag{3.27}$$

Set

$$\mathcal{P}^v_\mu := \inf_{\lambda\in\mathbb{R},\, \zeta\in\mathcal{M}_{0,1-q}} \mathcal{P}^v_\mu(\lambda,\zeta). \tag{3.28}$$

This will be the Parisi formula for the limit of $\mathbb{E}F^h_{N,1}(m)$ when $h_i = v(m_i)$ and the empirical measure μ_m in (3.23) converges weakly to μ.

THEOREM 3.2 (Parisi formula on the band). *Assume that $v \in C([0,1])$ and $\mu \in \mathcal{M}_{0,1}$. Let $m^N \in [0,1]^N$ be any sequence so that μ_{m^N} in* (3.23) *converges to μ weakly and let $h_i = v(m_i)$ for $i \le N$. Then*

$$\lim_{N\to\infty} \mathbb{E}F^h_{N,1}(m^N, \varepsilon_N) = \mathcal{P}^v_\mu, \tag{3.29}$$

provided that ε_N goes to zero slowly enough.

Remark 3.3. In Proposition 8.3 below, we will establish a connection between $\Phi_{a,\zeta}$ and Φ_ζ (recall (2.6)), which states that these PDE solutions are essentially the same up to a transformation. An important consequence of this connection is that the function $\Psi(a,\zeta)$ in (3.26) coincides with the one defined in (2.36) up to a shift. With a properly chosen external field (see (3.31) below), this allows us to simplify the above Parisi formula on the band and naturally gives rise to the desired TAP correction (defined in (2.24)).

3.4 Finding the effective external field

Next, in order to obtain (3.21), we aim to find an external field such that the model with the Hamiltonian $H^m_N(\widetilde{\boldsymbol{\sigma}}) + \sum_{i=1}^N h_i\widetilde{\sigma}_i$ on the band $B(m,\varepsilon)$ has zero in the support of its Parisi measure. Given a function $v \in C([0,1])$, suppose that (λ^*, ζ^*) is the minimizer in the Parisi formula (3.28). In the proof, we will deal with cases when $\lambda^* = +\infty$, but for the purpose of this discussion let us assume that it is finite. For zero to be in the support of ζ^*, it is necessary and sufficient that

$$\partial_x \Phi_{a,\zeta^*}\big(0, \lambda^* a + v(a)\big) = 0 \tag{3.30}$$

for all a in the support of μ. We will not prove this (standard) statement, because we will only need an implication in one direction (that will appear in Theorem 3.4(v) below), but, again, let us use it as a motivation for what we do next. If we include the term $\lambda^* a$ into the field $v(a)$, then the field must satisfy

$$\partial_x \Phi_{a,\zeta^*}(0, v(a)) = 0 \quad \text{for all } a$$

in the support of μ. If we recall the definition (3.26) above, this means that our only hope to force zero in the support of the Parisi measure is to restrict our attention to external fields generated by functions of the form

$$v_\zeta(a) := \Psi(a,\zeta) \tag{3.31}$$

for $\zeta \in \mathcal{M}_{0,1-q}$, and, moreover, with such choice of v_ζ, the Parisi measure ζ^* must coincide with ζ, so that (3.30) holds

$$\partial_x \Phi_{a,\zeta}(0, v_\zeta(a)) = 0. \tag{3.32}$$

Actually, this will automatically force $\lambda^* = 0$, so this equation matches (3.30). This raises two issues.

(1) First of all, we will see that $\Psi(a,\zeta)$ goes to $+\infty$ as a approaches 1 uniformly over ζ (see Lemma 13.3), while v in the Parisi formula in Theorem 3.2 was continuous on $[0,1]$, so we cannot apply the Parisi formula with $v = v_\zeta$ when $1 \in \mathrm{supp}(\mu)$.
(2) The minimizer ζ^* depends on the external field v_ζ, and we want it to coincide with ζ in the definition of v_ζ. How can we find such a 'good' choice of ζ?

The second issue can be solved via an implicit fixed point problem, using the Schauder fixed point theorem; however, we will give a more direct and explicit way to find such good ζ. The first issue will be handled by an approximation argument, which, at a crucial step, will allow us to work with measures μ with the support separated from 1. Because these two issues present very different obstacles, it will be convenient to work with two intermediate definitions of the functional $\mathrm{TAP}(\mu)$ in (2.25), which will be shown to coincide with it.

Let us consider the following growth condition on functions $v\colon [0,1] \to \mathbb{R}$,

$$v(a) \le c_1 + c_2 \tanh^{-1}(a) \ \text{ for } a \in [0,1), \tag{3.33}$$

for some absolute constants $c_1, c_2 > 0$ that depend only on the model ξ and can be found explicitly from the proof of Lemma 13.3 below. Set

$$\begin{aligned} V &= \big\{v \in C([0,1]) : v \ge 0, v \text{ is nondecreasing, and (3.33) holds}\big\}, \\ \overline{V} &= \big\{v \in C([0,1)) : v \ge 0, v \text{ is nondecreasing, and (3.33) holds}\big\}. \end{aligned} \tag{3.34}$$

Note that the only difference between V and $\overline{V}$ is on the right boundary of v. We will prove that the functions $v_\zeta(a) = \Psi(a,\zeta)$ belong to $\overline{V}$ (see Lemma 13.2), which is the real reason behind these definitions. Let us recall the functional $\mathcal{P}^v_\mu(\lambda,\zeta)$ defined in (3.27) for $v \in V$,

$$\mathcal{P}^v_\mu(\lambda,\zeta) := \int_{[0,1]} \Phi_{a,\zeta}(0, \lambda a + v(a)) d\mu(a) - \frac{1}{2}\int_0^{1-q} s\xi_q''(s)\zeta(s)ds,$$

and define a new functional $\overline{\mathcal{P}}^v_\mu(\lambda,\zeta)$ for $v \in \overline{V}$,

$$\overline{\mathcal{P}}^v_\mu(\lambda,\zeta) := \int_{[0,1)} \Phi_{a,\zeta}(0, \lambda a + v(a)) d\mu(a) - \frac{1}{2}\int_0^{1-q} s\xi_q''(s)\zeta(s)ds. \tag{3.35}$$

For $\mu \in \mathcal{M}_{0,1}$ such that $\int a^2 d\mu(a) = q$, we define

$$\begin{aligned} \mathbf{T}(\mu) &:= \inf_{v \in V,\, \zeta \in \mathcal{M}_{0,1-q}} \mathcal{P}_\mu^v(0,\zeta), \\ \overline{\mathbf{T}}(\mu) &:= \inf_{v \in \overline{V},\, \zeta \in \mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^v(0,\zeta). \end{aligned} \tag{3.36}$$

In the case that $\operatorname{supp}(\mu) \subset [0,1)$, it is evident that these coincide, and we will in fact show that they always coincide. The first representation in (3.36) will be convenient when working with the Parisi formula on the band, and the second representation will be convenient for analytical reasons and because it allows us to find a good choice of ζ directly. We summarize all these properties in the following theorem.

We define by

$$\theta_q \zeta(t) = \zeta(t+q) \quad \text{for } t \in [0, 1-q] \tag{3.37}$$

the shift operator $\theta_q \colon \mathcal{M}_{0,1} \to \mathcal{M}_{0,1-q}$.

THEOREM 3.4. *The following statements hold*:

(i) *If $\mu \in \mathcal{M}_{0,1}$ and $q = \int x^2 \, d\mu(a)$, then*

$$\mathrm{TAP}(\mu, \zeta') = \overline{\mathcal{P}}_\mu^{v_\zeta}(0,\zeta)$$

for all $\zeta' \in \mathcal{M}_{0,1}$ and $\zeta = \theta_q \zeta'$, and

$$\mathrm{TAP}(\mu) = \mathbf{T}(\mu) = \overline{\mathbf{T}}(\mu) = \inf_{\zeta \in \mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_\zeta}(0,\zeta). \tag{3.38}$$

(ii) *The functional* $\mathrm{TAP}(\mu)$ *is continuous on* $(\mathcal{M}_{0,1}, d_1)$.

(iii) *The right-hand side in* (3.38) *has a unique minimizer* ζ_0.

(iv) *The minimizer ζ_0 in (iii) satisfies*

$$\overline{\mathcal{P}}_\mu^{v_{\zeta_0}}(0,\zeta_0) = \inf_{\lambda \in \mathbb{R}, \zeta \in \mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_{\zeta_0}}(\lambda,\zeta) \quad (=: \overline{\mathcal{P}}_\mu^{v_{\zeta_0}}).$$

(v) *The minimizer ζ_0 in (iii) has zero in its support,* $0 \in \operatorname{supp}(\zeta_0)$.

(vi) *If there exists some $\zeta_1 \in \mathcal{M}_{0,1-q}$ such that $(0,\zeta_1)$ is a minimizer of*

$$\inf_{\lambda \in \mathbb{R}, \zeta \in \mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(\lambda,\zeta);$$

then $\zeta_0 = \zeta_1$.

Notice that if $\mu = \delta_1$ then $q = 1$, so $\overline{\mathcal{P}}_\mu^v(\lambda,\zeta) \equiv 0$, $\mathcal{M}_{0,1-q} = \{\delta_0\}$, and the claims (iii)–(vi) hold trivially. If $\operatorname{supp}(\mu) \subseteq \{0,1\}$, then $\overline{\mathcal{P}}_\mu^v(\lambda,\zeta)$ does not depend on λ. In Remark 13.5 below, we will see that in all other cases the minimizer $(\lambda_\mu^v, \zeta_\mu^v)$ exists and is unique. For convenience, we will use the convention that $(\lambda_\mu^v, \zeta_\mu^v) = (0, \delta_0)$ when $\mu = \delta_1$, and $(\lambda_\mu^v, \zeta_\mu^v) = (0, \zeta_\mu^v)$ when $\operatorname{supp}(\mu) \subseteq \{0,1\}$.

3.5 Classical TAP correction

Finally, let us explain how the generalized TAP correction defined above leads to the classical TAP correction and the Plefka condition in [53, 67]. In this section, we will consider the nontrivial case $\mu \neq \delta_1$ and will describe when the minimizer ζ_0 in Theorem 3.4 is *replica symmetric*, $\zeta_0 = \delta_0$. Let us denote the corresponding replica symmetric external field by

$$v_{\mathrm{RS}}(a) := v_{\delta_0}(a) = \Psi(a, \delta_0) \quad \forall a \in [0,1). \tag{3.39}$$

Using the Cole-Hopf transformation yields

$$\begin{aligned} \Phi_{a,\delta_0}(s,x) &= \log 2 + \log \mathbb{E} \cosh(a + t(s)g) e^{-a(x+t(s)g)} \\ &= \frac{1+a^2}{2} t^2(s) - ax + \log 2 \cosh(x - at^2(s)), \end{aligned} \tag{3.40}$$

where g is a standard Gaussian random variable and, for $s \in [0, 1-q]$, we denoted

$$t(s) := \big(\xi_q'(1-q) - \xi_q'(s)\big)^{1/2}.$$

From this formula, we can express v_{RS} explicitly as

$$v_{\mathrm{RS}}(a) := \tanh^{-1}(a) + a\xi_q'(1-q) \ \ \forall a \in [0,1). \tag{3.41}$$

By Theorem 3.4 (iv), if δ_0 is the minimizer of the functional in (2.25), we must have

$$\overline{\mathcal{P}}_\mu^{v_{\mathrm{RS}}}(0, \delta_0) = \inf_{\lambda, \zeta \in \mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_{\mathrm{RS}}}(\lambda, \zeta). \tag{3.42}$$

On the other hand, by Theorem 3.4 (vi), if this holds then δ_0 is the minimizer. Hence, in order to characterize the replica symmetric TAP correction, it is enough to describe when (3.42) holds. Recall (3.40) and define

$$\gamma_\mu(s) := \int_{[0,1)} \mathbb{E}\big[\partial_x \Phi_{a,\delta_0}(s, g_a(s))^2 \exp\big(\Phi_{a,\delta_0}(s, g_a(s)) - \Phi_{a,\delta_0}(0, g_a(0))\big)\big] \, d\mu(a),$$

where $g_a(s) = v_{\mathrm{RS}}(a) + \xi_q'(s)^{1/2} g$. For $s \in [0,1]$, define an auxiliary function Γ_μ by

$$\Gamma_\mu(s) = \int_0^s \xi_q''(r)(\gamma_\mu(r) - r) dr. \tag{3.43}$$

We derive this function through the directional derivative of the functional $\overline{\mathcal{P}}_\mu^{v_{\mathrm{RS}}}(0, \cdot)$ (see the derivation in (12.1) below). Recall $I(a)$ and $C(q)$ in (2.31). The following holds.

PROPOSITION 3.5. *Assume that* $\delta_1 \neq \mu \in \mathcal{M}_{0,1}$ *and* $q = \int a^2 \, d\mu(a)$. *Then* δ_0 *is the minimizer of* (2.25) (*i.e.,* (3.42) *holds*) *if and only if*

$$\Gamma_\mu(s) \leq 0 \ \ \forall s \in [0, 1-q]. \tag{3.44}$$

Furthermore, in this case, $\mathrm{TAP}(\mu)$ *is given by the classical TAP correction,*

$$\mathrm{TAP}(\mu) = -\int I(a)d\mu(a) + C(q). \tag{3.45}$$

The above proposition will be used to conclude Corollary 2.3 from the general TAP representation of Theorem 2.2. The proposition also naturally leads to the so-called Plefka condition in the SK model, that is, $\xi(s) = \beta^2 s^2/2$.

PROPOSITION 3.6. *Let* $\xi(s) = \beta^2 s^2/2$. *Assume that* $\delta_1 \neq \mu \in \mathcal{M}_{0,1}$. *If* δ_0 *is the minimizer of* (2.25), *i.e.,* (3.44) *holds, then the so-called* Plefka's condition *holds,*

$$\beta^2 \int (1-a^2)^2 \, d\mu(a) \leq 1. \tag{3.46}$$

Remark 3.7. If we consider the pure SK model with $\xi(s) = \beta^2 s^2/2$ with nonzero external field h, one can see that the generalized TAP correction is not always replica symmetric on the sphere $\frac{1}{N}\|m\|^2 = q_{\mathrm{EA}}$, where the Edwards-Anderson parameter q_{EA} is the largest point in the support of the Parisi measure. Let us now consider (β, h) below the AT line

$$\beta^2 \mathbb{E} \frac{2}{\cosh^4(\beta z\sqrt{q} + h)} \leq 1,$$

where q is the unique solution of $q = \mathbb{E}\tanh^2(\beta z\sqrt{q} + h)$. It is well-known based on simulations, and in fact proved in some region of parameters in [33], that below the AT line the original model is replica symmetric, and $q_{\mathrm{EA}} = q$. Let us now take m that has Nq coordinates equal to 1 and $N(1-q)$ coordinates equal to 0. For such m, Plefka's condition becomes

$$\beta^2 \int (1-a^2)^2 \, d\mu_m(a) = \beta^2(1-q) = \beta^2 \mathbb{E} \frac{1}{\cosh^2(\beta z\sqrt{q} + h)} \leq 1.$$

Clearly, we can choose (β, h) below the AT line such that Plefka's condition is violated, which means that the generalized TAP correction does not always coincide with the classical TAP correction on the sphere $\frac{1}{N}\|m\|^2 = q_{\mathrm{EA}}$.

In Plefka [53], it was conjectured that in the SK model,

$$\lim_{N\to\infty} F_N = \lim_{\varepsilon \downarrow 0} \lim_{N\to\infty} \sup\left(\frac{H_N(m)}{N} - \int I(a)d\mu_m(a) + C(q) \right),$$

where the supremum is taken over all $m \in [-1,1]^N$ satisfying that $q = \|m\|_2^2/N$ and the Plefka condition, $\beta^2 \int (1-a^2)^2 \, d\mu_m(a) \leq 1$. While Plefka obtained this condition through the consideration of the convergence criterion for the series expansion of the SK free energy, we discover the same condition from an analogous study of the so-called Almeida-Thouless line for the Parisi formula on the band with external field v_{RS}; namely, it is determined by the second derivative of Γ_μ at

$s = 0$. We anticipate that it is not always true that when the Plefka condition is satisfied, (3.44) is automatically valid. In addition, we mention that it looks possible that one can find a μ such that the Plefka condition is satisfied, but

$$\overline{\mathcal{P}}_{\mu}^{\upsilon_{\mathrm{RS}}}(0, \delta_0) = -\int I(a) d\mu(a) + C(q)$$

is strictly larger than $\mathrm{TAP}(\mu)$. It is, however, not clear to us how to compare the corresponding TAP free energy of such μ with our expression (2.24), so the validity of Plefka's conjecture remains unclear.

Let us describe the structure of the rest of the paper. In the next section we derive the uniform statement in Theorem 2.1, conditional on Theorem 3.4 and a representation of the limiting replicated free energy on a band at the level of expectation. Using Theorem 2.1, in Section 5 we prove the generalized TAP representation in Theorem 2.2. In Section 6, we reduce the TAP correction to the free energy of the replicated model on the narrow band, and in Section 7 we prove the Parisi formula for this model stated in Theorem 3.2. In Section 8 we prove the properties of various TAP representations in Theorem 3.4. Section 9 contains the key step, which combines the Parisi formula with the choice of the optimal external field from Theorem 3.4 to derive the first representation $\mathbf{T}(\mu)$ in (3.36) at the level of expectations. In Section 10, we prove Theorem 2.4, and we compute the gradient of $\mathrm{TAP}(\mu_m)$ in Theorem 2.5 in Section 11. In Section 12 we prove Corollary 2.3 and Propositions 3.5 and 3.6, about the classical TAP correction. Finally, in Section 13, we prove various technical results used throughout the paper.

4 Uniform TAP Correction

In this section, we will prove our main result in Theorem 2.1. Our proof is conditional on the continuity of the mapping $\mu \to \mathrm{TAP}(\mu)$ that we stated in Theorem 3.4(ii) above, which we will prove in Section 8, and the following two lemmas relating the limit of the expectations $\mathbb{E}\,\mathrm{TAP}_{N,n}(m, \varepsilon, \delta)$ to the functional $\mathrm{TAP}(\mu)$, which we will prove in Section 9.

LEMMA 4.1. *For any sequence $m = m^N \in [0,1]^N$ such that $\mu_m \to \mu \in \mathcal{M}_{0,1}$,*

$$\inf_{\varepsilon,\delta,n} \limsup_{N\to\infty} \mathbb{E}\,\mathrm{TAP}_{N,n}(m, \varepsilon, \delta) \le \mathrm{TAP}(\mu). \tag{4.1}$$

LEMMA 4.2. *Take any $\eta \in (0,1)$. For any sequence $m = m^N \in [0, 1-\eta]^N$ such that $\mu_m \to \mu \in \mathcal{M}_{0,1}$ and $\mathrm{supp}(\mu) \subseteq [0, 1-\eta]$,*

$$\inf_{\varepsilon,\delta,n} \liminf_{N\to\infty} \mathbb{E}\,\mathrm{TAP}_{N,n}(m, \varepsilon, \delta) \ge \mathrm{TAP}(\mu). \tag{4.2}$$

To move from the statements about expectations above to the actual random TAP free energies, we will need the following basic, but crucial, concentration result. This is another key way in which we utilize many orthogonal directions.

LEMMA 4.3. *For some constant $c_\xi > 0$ depending only on ξ,*

$$\mathbb{P}\{|\operatorname{TAP}_{N,n}(m,\varepsilon,\delta) - \mathbb{E}\operatorname{TAP}_{N,n}(m,\varepsilon,\delta)| > t\} < 2\exp\left(-\frac{Nt^2 c_\xi}{1/n+\delta+\varepsilon}\right). \tag{4.3}$$

PROOF. Note that, up to normalization by n, $\operatorname{TAP}_{N,n}(m,\varepsilon,\delta)$ is the free energy of the Gaussian process $\widetilde{H}_N(\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n) := \sum_{i=1}^n (H_N(\boldsymbol{\sigma}^i) - H_N(m))$ on $B_n(m,\varepsilon,\delta)$, which, using (3.1) and (2.12), has variance bounded by

$$\begin{aligned}\frac{1}{N}\mathbb{E}\big(\widetilde{H}_N(\boldsymbol{\sigma}^1,\dots,\boldsymbol{\sigma}^n)\big)^2 &\le 4n\xi(1) + n(n-1)\xi'(1)(\delta+2\varepsilon)\\ &\le c_\xi(n+n^2\delta+n^2\varepsilon).\end{aligned}$$

The lemma therefore follows from a canonical Gaussian concentration result; see, e.g., [44, theorem 1.2]. □

We will also need the following consequence of the above concentration.

LEMMA 4.4. *For any $k \in \mathbb{N}$ there exists $A_N \subseteq (0,1)^N$ with at most $e^{N\log k}$ elements such that letting $\eta = 1/k$, for every $m \in A_N$,*

$$\max_{1\le i\le N} m_i \le 1 - \frac{\eta}{2}, \tag{4.4}$$

and, for every $m \in [0,1]^N$, there exists $m' \in A_N$ such that

$$\max_{i\le N} |m_i - m_i'| \le \frac{\eta}{2}. \tag{4.5}$$

Furthermore, there exists a constant $c_\xi > 0$ such that for any $\varepsilon,\delta > 0$ and $t, c > 0$ satisfying

$$\frac{-t^2 c_\xi}{1/n+\delta+\varepsilon} + \log\frac{2}{\eta} < -c, \tag{4.6}$$

we have that

$$\mathbb{P}\left(\max_{m\in A_N}\left|\operatorname{TAP}_{N,n}(m,\varepsilon,\delta) - \mathbb{E}\operatorname{TAP}_{N,n}(m,\varepsilon,\delta)\right| > t\right) < 2e^{-Nc}. \tag{4.7}$$

PROOF. Let P be a regular partition of $[0,1]^N$ using cubes of edge length η, and define A_N as the collection of all center points of the cubes in P. Of course, (4.4) and (4.5) hold, and A_N has $(1/\eta)^N = e^{N\log k}$ elements. Finally, (4.7) follows from Lemma 4.3 by a union bound. □

4.1 Proof of the upper bound of (2.28)

Let t, $c > 0$ and assume towards the contradiction that, for some ε_N, $\delta_N \to 0$ and $n_N \to \infty$, on some subsequence in N,

$$\mathbb{P}\big(\exists m \in [-1,1]^N : \operatorname{TAP}_{N,n_N}(m,\varepsilon_N,\delta_N) \ge \operatorname{TAP}(\mu_m) + t\big) \ge e^{-Nc}. \tag{4.8}$$

Let us choose $\eta_N \downarrow 0$ in such a way that

$$\frac{-(t/4)^2 c_\xi}{1/n_N + \delta_N + \varepsilon_N + 3\eta_N} + \log \frac{2}{\eta_N} < -2c. \tag{4.9}$$

Let us take the set A_N in Lemma 4.4 so that

$$\begin{aligned}\mathbb{P}\Big(\max_{m \in A_N} \big| &\mathrm{TAP}_{N,n_N}(m, \varepsilon_N + 3\eta_N, \delta_N) \\ &- \mathbb{E}\,\mathrm{TAP}_{N,n_N}(m, \varepsilon_N + 3\eta_N, \delta_N) \big| > t/4 \Big) < 2e^{-2Nc}.\end{aligned} \tag{4.10}$$

Suppose that m' satisfies the inequality of (4.8), and let $m \in A_N$ be such that (4.5) holds. Since $\|m - m'\|_2 < \sqrt{N}\eta_N$, we have

$$B(m, \varepsilon_N + 3\eta_N) \supseteq B(m', \varepsilon_N)$$

and, moreover, $d_1(\mu_m, \mu_{m'}) < \eta_N$. From the (uniform) continuity of $\mu \to \mathrm{TAP}(\mu)$ in Theorem 3.4 (ii),

$$\mathrm{TAP}_{N,n_N}(m, \varepsilon_N + 3\eta_N, \delta_N) \geq \mathrm{TAP}(\mu_{m'}) + t \geq \mathrm{TAP}(\mu_m) + t/2$$

for large N and, on the subsequence as above,

$$\mathbb{P}\big(\exists m \in A_N : \mathrm{TAP}_{N,n_N}(m, \varepsilon_N + 3\eta_N, \delta_N) \geq \mathrm{TAP}(\mu_m) + t/2\big) \geq e^{-Nc}.$$

From (4.10), (deterministically)

$$\exists m \in A_N : \mathbb{E}\,\mathrm{TAP}_{N,n_N}(m, \varepsilon_N + 3\eta_N, \delta_N) \geq \mathrm{TAP}(\mu_m) + t/4. \tag{4.11}$$

Since $(\mathcal{M}_{-1,1}, d_1)$ is a compact space, there exists a subsequence $m = m^N \in A_N$ such that $\mu_{m^N} \to \mu$ for some $\mu \in \mathcal{M}_{-1,1}$ and (along this subsequence)

$$\limsup_{N \to \infty} \mathbb{E}\,\mathrm{TAP}_{N,n_N}(m, \varepsilon_N + 3\eta_N, \delta_N) \geq \mathrm{TAP}(\mu) + t/4. \tag{4.12}$$

This contradicts Lemma 4.1, which finishes the proof of the upper bound. □

4.2 Proof of the lower bound of (2.28)

The proof is a variation of that of the upper bound. Let $t, c > 0$ and assume towards contradiction that for some ε, δ as small as we wish and n as large as we wish, there exists a subsequence in N such that

$$\mathbb{P}\big(\exists m \in [-1,1]^N : \mathrm{TAP}_{N,n}(m, \varepsilon, \delta) \leq \mathrm{TAP}(\mu_m) - t\big) \geq e^{-Nc}. \tag{4.13}$$

Let us make sure that ε, δ, n are such that we can choose $\eta < \varepsilon/6$ satisfying

$$\frac{-(t/4)^2 c_\xi}{1/n + \delta + \varepsilon/2} + \log \frac{2}{\eta} < -2c,$$

and such that $d_1(\mu, \mu') \leq \eta$ implies that $|\mathrm{TAP}(\mu) - \mathrm{TAP}(\mu')| \leq t/2$. By Lemma 4.4, there exists A_N such that (4.4) and (4.5) hold and

$$\begin{aligned}\mathbb{P}\Big(\max_{m \in A_N} \big| &\mathrm{TAP}_{N,n}(m, \varepsilon/2, \delta) - \mathbb{E}\,\mathrm{TAP}_{N,n}(m, \varepsilon/2, \delta) \big| > t/4 \Big) \\ &< 2e^{-2Nc}.\end{aligned} \tag{4.14}$$

Suppose that m' satisfies the inequality of (4.13), and let $m \in A_N$ be such that (4.5) holds. Since $\|m - m'\|_2 < \sqrt{N}\eta$ and $\eta < \varepsilon/6$, we have

$$B(m', \varepsilon) \supseteq B(m', \varepsilon/2 + 3\eta) \supseteq B(m, \varepsilon/2),$$

and, moreover, $d_1(\mu_m, \mu_{m'}) < \eta$. Therefore,

$$\mathrm{TAP}_{N,n}(m, \varepsilon/2, \delta) \leq \mathrm{TAP}_{N,n}(m', \varepsilon, \delta) \leq \mathrm{TAP}(\mu_{m'}) - t \leq \mathrm{TAP}(\mu_m) - t/2.$$

This implies that, on the subsequence as in (4.13),

$$\mathbb{P}\left(\exists m \in A_N : \mathrm{TAP}_{N,n}(m, \varepsilon/2, \delta) \leq \mathrm{TAP}(\mu_m) - t/2\right) \geq e^{-Nc},$$

and, therefore, from (4.14), (deterministically)

$$\exists m \in A_N : \mathbb{E}\,\mathrm{TAP}_{N,n}(m, \varepsilon/2, \delta) \leq \mathrm{TAP}(\mu_m) - t/4. \tag{4.15}$$

Since $(\mathcal{M}_{-1,1}, d_1)$ is a compact space, there exists a subsequence $m = m^N \in A_N$ such that $\mu_{m^N} \to \mu$ for some $\mu \in \mathcal{M}_{-1,1}$ and (along this subsequence)

$$\liminf_{N\to\infty} \mathbb{E}\,\mathrm{TAP}_{N,n}(m, \varepsilon/2, \delta) \leq \mathrm{TAP}(\mu) - t/4. \tag{4.16}$$

However, the condition (4.4) implies that $\mathrm{supp}(\mu) \subseteq [0, 1 - \eta/2]$, and the above inequality contradicts Lemma 4.2. This finishes the proof. □

5 Generalized TAP representation

In this section, we prove the generalized TAP representation in Theorem 2.2 using the concentration of Theorem 2.1. The basic idea is similar to the proof of [59, Lemma 17]. We note that our proof of Theorem 2.1 in Section 4 is conditional on the results stated in the beginning of that section, and thus so is the current proof.

PROOF. Let $\varepsilon, \delta > 0$ and $n \geq 1$. By definition, for any m,

$$F_N \geq \frac{H_N(m)}{N} + \mathrm{TAP}_{N,1}(m, \varepsilon, \delta) \geq \frac{H_N(m)}{N} + \mathrm{TAP}_{N,n}(m, \varepsilon, \delta).$$

In particular, from the uniform convergence of Theorem 2.1,

$$\lim_{N\to\infty} \mathbb{P}\left(F_N > \max_{\frac{\|m\|^2}{N}=q} \left(\frac{H_N(m)}{N} + \mathrm{TAP}(\mu_m) \right) - t \right) = 1.$$

Fix some value q that belongs to the support of the Parisi measure of the model (1.2). It is well-known that the free energy F_N concentrates at exponential rate around its mean for large N, see, e.g., [44, theorem 1.2]. By the Borell-TIS inequality so does the maximum in (2.29). Combining the above with Theorem 2.1,

we conclude that to complete the proof of Theorem 2.2 it will be enough to show that for any small $t, \varepsilon, \delta > 0$ and large $n \geq 1$,

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}\Big(\exists m \in [-1,1]^N,\ \frac{1}{N}\|m\|^2 = q : F_N < \frac{H_N(m)}{N} + \mathrm{TAP}_{N,n}(m,\varepsilon,\delta) + t \Big) = 0. \tag{5.1}$$

>From [48, lemma 4.8], for any $\tau, t > 0$ and $n \geq 1$ we have that

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}\left(\frac{1}{N} \log G_N^{\otimes 2n}\big\{ \forall i \neq j,\ \big|R(\boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j) - q\big| < \tau \big\} > -t \right) = 0. \tag{5.2}$$

By conditioning on $\boldsymbol{\sigma}^{n+1}, \dots, \boldsymbol{\sigma}^{2n}$, we conclude that, with probability not exponentially small in N (w.r.t. the disorder only) there exist $\overline{\boldsymbol{\sigma}}^{n+1}, \dots, \overline{\boldsymbol{\sigma}}^{2n} \in \Sigma_N$ such that for any $n+1 \leq i \neq j \leq 2n$, $|R(\overline{\boldsymbol{\sigma}}^i, \overline{\boldsymbol{\sigma}}^j) - q| < \tau$, and such that

$$\begin{aligned} &\frac{1}{nN} \log G_N^{\otimes 2n}\Big\{ \forall i \neq j \leq 2n,\ \big|R(\boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j) - q\big| < \tau \ \Big|\ \boldsymbol{\sigma}^{n+i} = \overline{\boldsymbol{\sigma}}^{n+i},\ i = 1, \dots, n \Big\} \\ &\quad = \frac{1}{nN} \log \sum_{A_n(\tau)} \exp\Big\{ \sum_{i=1}^n H_N(\boldsymbol{\sigma}^i) \Big\} - F_N > -t, \end{aligned} \tag{5.3}$$

where we define $A_n(\tau)$ as the set of points $(\boldsymbol{\sigma}^1, \dots, \boldsymbol{\sigma}^n) \in \Sigma_N^n$ such that

$$|R(\boldsymbol{\sigma}^i, \boldsymbol{\sigma}^j) - q| < \tau \quad \text{and} \quad |R(\boldsymbol{\sigma}^i, \overline{\boldsymbol{\sigma}}^k) - q| < \tau,$$

for any $i, j \leq n$ and $n+1 \leq k \leq 2n$.

Define $\overline{\boldsymbol{\sigma}} := n^{-1} \sum_{i=n+1}^{2n} \overline{\boldsymbol{\sigma}}^i$ and set $m = 0$ if $\overline{\boldsymbol{\sigma}} = 0$, and $m = \sqrt{Nq}\,\overline{\boldsymbol{\sigma}}/\|\overline{\boldsymbol{\sigma}}\|$ otherwise. Given $\delta, \varepsilon > 0$, it is straightforward to check that $A_n(\tau) \subset B_n(m, \varepsilon, \delta)$, provided that $\tau > 0$ is small enough and $n \geq 1$ is large enough. Assuming this inclusion and assuming that (5.3) holds, m is a point as in (5.1), since

$$\frac{H_N(m)}{N} + \mathrm{TAP}_{N,n}(m,\varepsilon,\delta) = \frac{1}{nN} \log \sum_{B_n(m,\varepsilon,\delta)} \exp\Big\{ \sum_{i=1}^n H_N(\boldsymbol{\sigma}^i) \Big\}.$$

This completes the proof. □

6 Reduction to a Model on the Band

In this section, we will justify approximations of $\mathrm{TAP}_{N,n}(m,\varepsilon,\delta)$ by $F_{N,n}(m,\varepsilon,\delta)$. We will need the following technical lemma, which follows from a more general result in [10, cor. 59], but which we prove here for convenience.

LEMMA 6.1. *For a mixed p-spin model in* (1.2),

$$\mathbb{E} \max_{\|\boldsymbol{\sigma}\| \leq \sqrt{N}} \|\nabla H_N(\boldsymbol{\sigma})\| \leq c_\xi \sqrt{N}, \tag{6.1}$$

where one can take $c_\xi = 2(\sum_{p\geq 1} \beta_p^2 p^3)^{1/2}$.

PROOF. Let us represent the maximum above as

$$\max_{\|\boldsymbol{\sigma}\|\le\sqrt{N}} \|\nabla H_N(\boldsymbol{\sigma})\| = \max_{\|u\|\le 1} \max_{\|\boldsymbol{\sigma}\|\le\sqrt{N}} u\cdot\nabla H_N(\boldsymbol{\sigma}) \tag{6.2}$$

and denote

$$\begin{aligned} G(u,\boldsymbol{\sigma}) &:= u\cdot\nabla H_N(\boldsymbol{\sigma}) \\ &= \sum_{p\ge 1}\frac{\beta_p}{N^{(p-1)/2}}\sum_{i_1,\dots,i_p} g_{i_1,\dots,i_p}(u_{i_1}\cdots\sigma_{i_p}+\cdots+\sigma_{i_1}\cdots u_{i_p}). \end{aligned}$$

Using that $(a_1+\cdots+a_p)^2 \le p(a_1^2+\cdots+a_p^2)$ and the fact that $\|u\|\le 1$ and $\|\boldsymbol{\sigma}^j\|^2\le N$, one can easily check that

$$\mathbb{E}\big(G(u^1,\boldsymbol{\sigma}^1)-G(u^2,\boldsymbol{\sigma}^2)\big)^2 \le c_\xi^2\big(\|u^1-u^2\|^2+N^{-1}\|\boldsymbol{\sigma}^1-\boldsymbol{\sigma}^2\|\big),$$

where $c_\xi^2=\sum_{p\ge1}\beta_p^2p^3<\infty$. If we define, for i.i.d. standard Gaussian g_i and g_i',

$$\widetilde{G}(u,\boldsymbol{\sigma}) := c_\xi\bigg(\sum_{i=1}^N g_iu_i+N^{-1/2}\sum_{i=1}^N g_i'\sigma_i\bigg),$$

then the right-hand side above equals $\mathbb{E}(\widetilde{G}(u^1,\boldsymbol{\sigma}^1)-\widetilde{G}(u^2,\boldsymbol{\sigma}^2))^2$. By the Sudakov-Fernique inequality,

$$\mathbb{E}\max_{u,\boldsymbol{\sigma}} G(u,\boldsymbol{\sigma}) \le \mathbb{E}\max_{u,\boldsymbol{\sigma}}\widetilde{G}(u,\boldsymbol{\sigma}) \le 2c_\xi\sqrt{N},$$

and this finishes the proof. □

Let us start by proving the approximation in the equation (3.8).

LEMMA 6.2. *If* $q=R(m,m)=\|m\|^2/N>0$, *the equation* (3.8) *holds,*

$$\mathbb{E}\,\mathrm{TAP}_{N,n}(m,\varepsilon,\delta) = \frac{1}{nN}\mathbb{E}\log\sum_{B_n(m,\varepsilon,\delta)} e^{\sum_{i=1}^n \widehat{H}_N^m(\widetilde{\boldsymbol{\sigma}}^i)} + O\bigg(\frac{\varepsilon}{\sqrt{q}}\bigg), \tag{6.3}$$

with an implicit constant in the error term that depends only on ξ.

PROOF. Given $\boldsymbol{\sigma}$ and $\widetilde{\boldsymbol{\sigma}}=\boldsymbol{\sigma}-m$, define

$$\widetilde{\boldsymbol{b}} = \widetilde{\boldsymbol{\sigma}}-\frac{\widetilde{\boldsymbol{\sigma}}\cdot m}{m\cdot m}m, \quad \boldsymbol{b}=\boldsymbol{\sigma}-\frac{\widetilde{\boldsymbol{\sigma}}\cdot m}{m\cdot m}m=\widetilde{\boldsymbol{\sigma}}-\frac{\widetilde{\boldsymbol{\sigma}}\cdot m}{m\cdot m}m+m.$$

Notice that $\boldsymbol{b}$ is the projection of $\boldsymbol{\sigma}$ onto the hyperplane perpendicular to m passing through m, and $\widetilde{\boldsymbol{b}}=\boldsymbol{b}-m$. We view $\boldsymbol{b}$ as a function of $\boldsymbol{\sigma}$ but, to simplify notation, we will keep the function implicit. The projections satisfy $R(\boldsymbol{b}^1,\boldsymbol{b}^2)=R(\widetilde{\boldsymbol{b}}^1,\widetilde{\boldsymbol{b}}^2)+q$ and the Hamiltonians $H_N(\boldsymbol{b})-H_N(m)$ and $\widehat{H}_N^m(\widetilde{\boldsymbol{b}})$ are equal in distribution. As a result,

$$\frac{1}{nN}\mathbb{E}\log\sum_{B_n(m,\varepsilon,\delta)} e^{\sum_{\ell=1}^n\left[H_N(\boldsymbol{b}^\ell)-H_N(m)\right]} = \frac{1}{nN}\mathbb{E}\log\sum_{B_n(m,\varepsilon,\delta)} e^{\sum_{\ell=1}^n\widehat{H}_N^m(\widetilde{\boldsymbol{b}}^\ell)}.$$

Since, for $\boldsymbol{\sigma} \in B(m, \varepsilon)$,

$$\|\boldsymbol{\sigma} - \boldsymbol{b}\| = \|\widetilde{\boldsymbol{\sigma}} - \widetilde{\boldsymbol{b}}\| = \frac{\|\boldsymbol{\sigma} \cdot m - m \cdot m\|}{\|m\|} < \frac{\varepsilon}{\sqrt{q}}\sqrt{N},$$

in order to prove (3.8), it is enough to use that, by the above lemma,

$$\mathbb{E} \max_{\|\boldsymbol{\sigma}^1 - \boldsymbol{\sigma}^2\| \le \varepsilon\sqrt{N/q}} \left| H_N(\boldsymbol{\sigma}^1) - H_N(\boldsymbol{\sigma}^2) \right| \le c_\xi(\varepsilon/\sqrt{q})N, \tag{6.4}$$

and that a similar statement holds for $\widehat{H}_N^m$, which can be proved in exactly the same way. □

As we explained in the introduction, this implies that

$$\mathbb{E}\,\mathrm{TAP}_{N,n}(m, \varepsilon, \delta) = \mathbb{E} F_{N,n}(m, \varepsilon, \delta) + O\left(\frac{\varepsilon}{\sqrt{q}} + \left(\frac{1}{n} + \delta \right)^{1/2} \right). \tag{6.5}$$

When $q = R(m, m) = \|m\|^2/N$ is small and the above approximation is not good enough, let us say when $q < \varepsilon/2$, we will use a more straightforward reduction. We will not add and subtract the term $H_N(m)$, which in this case is small (on the scale $1/N$). Instead, the only modification we will make is to remove the external field in the original model. We will consider the Hamiltonian

$$H_N^0(\boldsymbol{\sigma}) = \sum_{p \ge 2} \beta_p H_{N,p}(\boldsymbol{\sigma}), \tag{6.6}$$

which is the original Hamiltonian in (1.2) with external field removed, so that

$$\mathbb{E} H_N^0(\boldsymbol{\sigma}^1) H_N^0(\boldsymbol{\sigma}^2) = N\xi_0\big(R(\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2)\big), \tag{6.7}$$

where $\xi_0(x) = \xi(x) - \xi'(0)x$ is defined exactly as $\xi_q(x)$ in (3.9) for $q = 0$. If we define

$$F_{N,n}^*(m, \varepsilon, \delta) := \frac{1}{nN} \log \sum_{B_n(m,\varepsilon,\delta)} e^{\sum_{\ell=1}^n H_N^0(\boldsymbol{\sigma}^\ell)}, \tag{6.8}$$

then the argument leading to (3.11) also gives in this case that

$$\mathbb{E}\,\mathrm{TAP}_{N,n}(m, \varepsilon, \delta) = \mathbb{E} F_{N,n}^*(m, \varepsilon, \delta) + O\left(\left(\frac{1}{n} + \delta \right)^{1/2} \right). \tag{6.9}$$

In this case of small q, the analogue of the Parisi formula in Theorem 3.2 is the following.

THEOREM 6.3. *Let $m^N \in [-1, 1]^N$ be any sequence such that $\mu_{m^N} \to \delta_0$. Then*

$$\lim_{N\to\infty} \mathbb{E} F_{N,1}^*(m^N, \varepsilon_N) = \inf_{\zeta \in \mathcal{M}_{0,1}} \left(\Phi_{0,\zeta}(0, 0) - \frac{1}{2} \int_0^1 s\xi_0''(s)\zeta(s)ds \right). \tag{6.10}$$

provided that ε_N goes to zero slowly enough.

This is, essentially, a classical Parisi formula for the original model without external field, only now we have the constraint $\boldsymbol{\sigma} \in B(m, \varepsilon_N)$. If $q = \|m\|^2/N \leq \varepsilon/2$, then, denoting $u = m/\|m\|$,

$$B(m, \varepsilon) \supseteq \{\boldsymbol{\sigma} : |\boldsymbol{\sigma} \cdot u| \leq \sqrt{\varepsilon/2}\},$$

and the last set is a constraint on the magnetization in some direction u. Since the Hamiltonian H_N^0 does not contain an external field, one can show that this constraint has no free energy cost, which explains why the above formula coincides with the Parisi formula for the unconstrained model. We are not going to give a proof of this for two reasons. First reason is that the proof is straightforward. The second reason is that the proof of Theorem 3.2 in the next section does this in a more complicated case, and following the same argument in this case would only significantly simplify the details.

7 Proof of the Parisi Formula on the Band

We establish the proof of Theorem 3.2 in this section. The argument is essentially the same as the treatment for the classical Sherrington-Kirkpatrick model, by utilizing Guerra's replica symmetry breaking scheme and the Aizenman-Sims-Starr scheme as implemented, for example, in [44]. The added complication here is that the spin configurations are re-centered and the external field varies for each site, depending on the function $v \in C([0, 1])$. These will require extra care of the uniform convergence of the free energy in the variable m. To this end, in Section 7.1, we first give a version of Theorem 3.2 in Proposition 7.1 in terms of the Ruelle probability cascades followed by a set of lemmas that are devoted to establishing uniform controls of various functionals. Sections 7.2 and 7.3 establish the upper and lower bounds in Proposition 7.1.

7.1 Ruelle probability cascades and continuity results

Let $q \in [0, 1)$. Denote by $\mathcal{M}^d_{0,1-q}$ the collection of all atomic $\zeta \in \mathcal{M}_{0,1-q}$ satisfying that, for some integer $r \geq 1$ and some sequences

$$0 < \zeta_0 < \cdots < \zeta_{r-1} < \zeta_r = 1, \tag{7.1}$$

$$0 = q_0 < q_1 < \cdots < q_r = 1 - q, \tag{7.2}$$

we have

$$\zeta(s) = \zeta_p \text{ if } q_p \leq s < q_{p+1} \text{ for some } 0 \leq p \leq r - 1. \tag{7.3}$$

Let $(v_\alpha)_{\alpha \in \mathbb{N}^r}$ be the weights of the Ruelle probability cascade [55] corresponding to the sequence (7.1) (see, e.g., section 2.3 in [44] for the definition). For $\alpha^1, \alpha^2 \in \mathbb{N}^r$, denote

$$\alpha^1 \wedge \alpha^2 = \max\{0 \leq p \leq r : \alpha^1_1 = \alpha^2_1, \ldots, \alpha^1_p = \alpha^2_p\}.$$

Let $f(x)$ be a function of the form $\sum_{p \geq 2} c_p^2 x^p$ such that $f(x_0) < \infty$ for some $x_0 > 1$. This ensures that all derivatives of f are well-defined on $(-x_0, x_0)$. The

main choice of f we have in mind is ξ_q, but it is convenient to keep it as a parameter in the following definitions. Let $\theta_f(x) = xf'(x) - f(x)$. Let $g_{f'}(\alpha)$ be a centered Gaussian process on $\mathbb{N}^r$ with the covariance given by

$$\mathbb{E} g_{f'}(\alpha^1) g_{f'}(\alpha^2) = f'(q_{\alpha^1 \wedge \alpha^2}), \tag{7.4}$$

and define $g_{\theta_f}(\alpha)$ similarly. For $m \in [0,1]^N$ and $\zeta \in \mathcal{M}^d_{0,1-q}$, we set

$$\Psi_N(m, \varepsilon, f, \zeta) = \frac{1}{N} \mathbb{E} \log \sum_{\alpha \in \mathbb{N}^r} v_\alpha \sum_{\sigma \in B(m,\varepsilon)} \exp \sum_{i \le N} \big(g_{f',i}(\alpha) + v(m_i) \big) \widetilde{\sigma}_i, \tag{7.5}$$

where $\widetilde{\sigma} = \sigma - m$ and $g_{f',i}$ are i.i.d. copies of $g_{f'}$. For $\zeta \in \mathcal{M}^d_{0,1-q}$, also set

$$\begin{aligned} \Upsilon(f, \zeta) &= \frac{1}{N} \mathbb{E} \log \sum_{\alpha \in \mathbb{N}^r} v_\alpha \exp \sqrt{N} g_{\theta_f}(\alpha) \\ &= \mathbb{E} \log \sum_{\alpha \in \mathbb{N}^r} v_\alpha \exp g_{\theta_f}(\alpha) = \frac{1}{2} \int_0^{1-q} \zeta(s) s f''(s) ds. \end{aligned} \tag{7.6}$$

The second and third equalities in this equation are well-known; see, e.g., [44, eq. (2.60)]. For $\zeta \in \mathcal{M}^d_{0,1-q}$ and $f = \xi_q$, we will denote

$$\Psi_N(m, \varepsilon, \zeta) := \Psi_N(m, \varepsilon, \xi_q, \zeta) \tag{7.7}$$

and

$$\Upsilon(\zeta) := \Upsilon(\xi_q, \zeta). \tag{7.8}$$

Recall that

$$F^h_{N,1}(m, \varepsilon) = \frac{1}{N} \log \sum_{\sigma \in B(m,\varepsilon)} e^{H^m_N(\widetilde{\sigma}) + \sum_{i=1}^N h_i \widetilde{\sigma}_i}. \tag{7.9}$$

Throughout the section we will assume that

$$h_i = v(m_i) \tag{7.10}$$

for some $v \in C([0,1])$ fixed. Recall the notation $\mu_m = \frac{1}{N} \sum_{i=1}^N \delta_{m_i}$.

PROPOSITION 7.1. *Given $\mu \in \mathcal{M}_{0,1}$ and an arbitrary sequence $m = m^N \in [0,1]^N$ such that $\mu_{m^N} \to \mu$ weakly if $q = \int x^2 \, d\mu(x)$, then*

$$\lim_{\varepsilon \downarrow 0} \lim_{N \to \infty} \mathbb{E} F^h_{N,1}(m^N, \varepsilon) = \inf_{\zeta \in \mathcal{M}^d_{0,1-q}} \big(\Psi(\mu, \zeta) - \Upsilon(\zeta) \big), \tag{7.11}$$

where the limit

$$\Psi(\mu, \zeta) := \lim_{\varepsilon \downarrow 0} \lim_{N \to \infty} \Psi_N(m^N, \varepsilon, \zeta) \tag{7.12}$$

exists and does not depend on the choice of the sequence (m^N).

More precisely, in (7.11) we mean that

$$\lim_{\varepsilon\downarrow 0}\liminf_{N\to\infty}\mathbb{E}\,F^h_{N,1}(m^N,\varepsilon)=\lim_{\varepsilon\downarrow 0}\limsup_{N\to\infty}\mathbb{E}\,F^h_{N,1}(m^N,\varepsilon),$$

and both are given by the right-hand side of (7.11). In particular, we can choose $\varepsilon_N\to 0$ slowly enough so that $\mathbb{E}\,F^h_{N,1}(m^N,\varepsilon_N)$ converges to the same limit.

Also, it is important to note that in $\Psi_N(m,\varepsilon,\zeta)$ the variables m and q are two independent parameters, and we will show that the quantity $\Psi(\mu,\zeta)$ in (7.12) is well-defined for any $\mu\in\mathcal{M}_{0,1}$, not necessarily satisfying the constraint $q=\int x^2\,d\mu(x)$. On the other hand, this constraint is crucial in (7.11), and that is why we minimize over $\zeta\in\mathcal{M}^d_{0,1-q}$.

For the rest of this subsection, we establish the convergence in (7.12), while the proof of (7.11) is deferred to the next two subsections. We begin with some basic continuity properties of the functionals defined above.

LEMMA 7.2. *For $\zeta\in\mathcal{M}^d_{0,1-q}$ and $\zeta'\in\mathcal{M}^d_{0,1-q'}$,*

$$|\Upsilon(f,\zeta)-\Upsilon(f,\zeta')|\le\frac{1}{2}f''(1)\bigg(\int_0^1|\zeta(s)-\zeta'(s)|ds+2|q-q'|\bigg). \tag{7.13}$$

and

$$\begin{aligned}&|\Psi_N(m,\varepsilon,f,\zeta)-\Psi_N(m,\varepsilon,f,\zeta')|\\&\quad\le 2f''(1)\bigg(\int_0^1|\zeta(s)-\zeta'(s)|ds+2|q-q'|\bigg).\end{aligned} \tag{7.14}$$

For any f_1 and f_2 as above and $\zeta\in\mathcal{M}^d_{0,1-q}$,

$$|\Psi_N(m,\varepsilon,f_1,\zeta)-\Psi_N(m,\varepsilon,f_2,\zeta)|\le\|f_1'-f_2'\|_\infty. \tag{7.15}$$

The last sup-norm is defined on $[0,1]$.

PROOF. The first inequality is clear. The proof of the second inequality is almost identical to the one in the classical SK model (due to Guerra [28]) and will be omitted. We only mention that the factor 2 instead of the usual $1/2$ is due to the fact that $|\widetilde{\sigma}|\le 2$ instead of 1 and we have the additional term $2|q-q'|$, because ζ and ζ' are defined on different intervals. If $q'<q$, we first need to interpolate all the parameters $q'_p>1-q$ down to $1-q$ before we start the usual argument, and one can check that all error along this interpolation is controlled by $4f''(1)|q-q'|$.

To prove (7.15), we replace the terms $g_{f_1',i}(\alpha)$ and $g_{f_2',i}(\alpha)$ by the usual Gaussian interpolation between the two, $\sqrt{t}\,g_{f_1',i}(\alpha)+\sqrt{1-t}\,g_{f_1',i}(\alpha)$. The derivative along the interpolation will be controlled by the maximum of $\big|f_1'\big(q_{\alpha^1\wedge\alpha^2}\big)-f_2'\big(q_{\alpha^1\wedge\alpha^2}\big)\big|$, and this finishes the proof. □

One immediate consequence of the above continuity properties that will be useful is the following.

LEMMA 7.3. *If* $\zeta \in \mathcal{M}^d_{0,1-q}$ *and* $\zeta' \in \mathcal{M}^d_{0,1-q'}$ *are such that* $\int_0^1 |\zeta(s)-\zeta'(s)|ds \le |q-q'|$ *then, for any* f_1, f_2,

$$|\Psi_N(m,\varepsilon,f_1,\zeta) - \Psi_N(m,\varepsilon,f_2,\zeta')| \le 6f_1''(1)|q-q'| + \|f_1' - f_2'\|_\infty. \tag{7.16}$$

PROOF. This follows from Lemma 7.2. □

Next, we will study continuity properties with respect to m. Recall the metric d_1 defined in (2.18) and note that if the coordinates of $m, m' \in [0,1]^N$ are arranged in the nondecreasing order, then

$$d_1(\mu_m, \mu_{m'}) = \frac{1}{N}\sum_{i=1}^N |m_i - m_i'| = \frac{1}{N}\|m-m'\|_1. \tag{7.17}$$

The following observation will be convenient.

LEMMA 7.4. *If* $\mu, \mu' \in \mathcal{M}_{0,1}$, *then, for* $k \ge 1$,

$$\left|\int x^k\, d\mu(x) - \int x^k\, d\mu'(x)\right| \le k d_1(\mu,\mu'). \tag{7.18}$$

PROOF. If μ^{-1} denotes the quantile transform of μ, then

$$\begin{aligned}
&\left|\int x^k d\mu(x) - \int x^k d\mu'(x)\right| \\
&\quad = \left|\int \mu^{-1}(x)^k dx - \int \mu'^{-1}(x)^k dx\right| \\
&\quad \le k\int \left|\mu^{-1}(x) - \mu'^{-1}(x)\right| dx = k\int |\mu(x)-\mu'(x)|dx,
\end{aligned}$$

which finishes the proof. □

For example, this implies that, for $m \in [0,1]^N$ and $\mu \in \mathcal{M}_{0,1}$,

$$\left|\frac{1}{N}\|m\|_2^2 - \int x^2\, d\mu(x)\right| = \left|\int x^2\, d\mu_m(x) - \int x^2\, d\mu(x)\right| \le 2d_1(\mu_m,\mu).$$

If $\boldsymbol{\sigma} \in B(m,\varepsilon)$, then the self-overlap of the recentered configuration $\widetilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - m$ can be rewritten as

$$\frac{1}{N}\widetilde{\boldsymbol{\sigma}}\cdot\widetilde{\boldsymbol{\sigma}} = 1 - \frac{1}{N}m\cdot m - \frac{2}{N}(\boldsymbol{\sigma}-m)\cdot m$$

and, therefore, if $q = \int x^2\, d\mu(x)$,

$$\left|\frac{1}{N}\widetilde{\boldsymbol{\sigma}}\cdot\widetilde{\boldsymbol{\sigma}} - (1-q)\right| \le 2(\varepsilon + d_1(\mu_m,\mu)). \tag{7.19}$$

In particular, if $\mu_m \to \mu$, then the self-overlap on the narrow band is approximately $1-q$, and the choice of $q_r = 1-q$ in (7.2) is designed to match this.

Let us introduce one more notation. For $v \in C([0,1])$, define

$$\Delta_v(x) = \inf_{t>0}\Big(\max_{|a-b|\le t} |v(a)-v(b)| + 2\|v\|_\infty \frac{x}{t}\Big). \tag{7.20}$$

By the uniform continuity of v, $\lim_{x\downarrow 0}\Delta_v(x) = 0$. We will use this quantity to control

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^N |v(m_i) - v(m_i')| \\
&\quad\le \max_{|a-b|\le t} |v(a)-v(b)| + 2\|v\|_\infty \frac{1}{N}\sum_{i=1}^N I(|m_i - m_i|' > t) \\
&\quad\le \max_{|a-b|\le t} |v(a)-v(b)| + 2\|v\|_\infty \frac{1}{t}\frac{1}{N}\sum_{i=1}^N |m_i - m_i|' \\
&\quad= \max_{|a-b|\le t} |v(a)-v(b)| + 2\|v\|_\infty \frac{1}{t}\, d_1(\mu_m, \mu_{m'}),
\end{aligned}$$

which implies that

$$\frac{1}{N}\sum_{i=1}^N |v(m_i) - v(m_i')| \le \Delta_v(d_1(\mu_m, \mu_{m'})). \tag{7.21}$$

The next lemma contains a key result that will later allow us to approximate general vectors m by nice 'discrete' ones.

LEMMA 7.5. *For $m, m' \in [0,1]^N$, if $N\varepsilon > 1$ and $d_1(\mu_m, \mu_{m'}) < \varepsilon^2/6$, then*

$$\big|\Psi_N(m,\varepsilon,f,\zeta) - \Psi_N(m',\varepsilon,f,\zeta)\big| < \Delta_f\big(d_1(\mu_m,\mu_{m'}),\varepsilon\big) \tag{7.22}$$

and

$$\big|\mathbb{E}F^h_{N,1}(m,\varepsilon) - \mathbb{E}F^h_{N,1}(m',\varepsilon)\big| < \Delta_\xi\big(d_1(\mu_m,\mu_{m'}),\varepsilon\big), \tag{7.23}$$

where

$$\Delta_f(d,\varepsilon) := 2\Delta_v(d) + \big(c_f + \|v\|_\infty\big)2d\left(1+\frac{3}{\varepsilon}\right) + \frac{3d}{\varepsilon}\log\frac{\varepsilon}{3d} \tag{7.24}$$

for some constant c_f that depends only on f.

Note that $v \in C([0,1])$, which implies that $\lim_{d\downarrow 0}\Delta_v(d) = 0$, and thus $\lim_{d\downarrow 0}\Delta_f(d,\varepsilon) = 0$ for any $\varepsilon > 0$. Let us also clarify that we will use c_ξ to denote various constants that depend on ξ_q for $q \in [0,1]$. The reason for this is that $\xi_q = \xi(x+q) - \xi(q) - \xi'(q)x$, and the derivatives of ξ_q can be controlled in terms of derivatives of ξ uniformly over q.

PROOF. Since the order of the coordinates of m does not affect Ψ_N, we can assume that they are arranged in the non-decreasing order. By (7.17) and the triangle inequality, for $\boldsymbol{\sigma} \in B(m', \varepsilon) \setminus B(m, \varepsilon)$,

$$\varepsilon < \frac{1}{N}|\boldsymbol{\sigma} \cdot m - m \cdot m| \le \frac{1}{N}|\boldsymbol{\sigma} \cdot m' - m' \cdot m'| + \frac{3}{N}\|m - m'\|_1 < \varepsilon + 3d_1(\mu_m, \mu_{m'}). \tag{7.25}$$

Since

$$\sum_{\sigma_i=1} m_i + \sum_{\sigma_i=-1} m_i = \|m\|_1, \quad \sum_{\sigma_i=1} m_i - \sum_{\sigma_i=-1} m_i = m \cdot \boldsymbol{\sigma},$$

we have that

$$\sum_{\sigma_i=1} m_i = (\|m\|_1 + m \cdot \boldsymbol{\sigma})/2, \quad \sum_{\sigma_i=-1} m_i = (\|m\|_1 - m \cdot \boldsymbol{\sigma})/2.$$

By (7.25), if $\boldsymbol{\sigma} \in B(m', \varepsilon) \setminus B(m, \varepsilon)$, then there are two possibilities:

$$N\varepsilon < \boldsymbol{\sigma} \cdot m - m \cdot m < N\big(\varepsilon + 3d_1(\mu_m, \mu_{m'})\big), \tag{7.26}$$

$$N\varepsilon < m \cdot m - \boldsymbol{\sigma} \cdot m < N\big(\varepsilon + 3d_1(\mu_m, \mu_{m'})\big). \tag{7.27}$$

In the first case,

$$\sum_{\sigma_i=1} m_i = (\|m\|_1 + m \cdot \boldsymbol{\sigma})/2 > (\|m\|_1 + \|m\|_2^2 + N\varepsilon)/2 \ge N\varepsilon/2,$$

and, in the second case, since $\|m\|_1 \ge \|m\|_2^2$,

$$\sum_{\sigma_i=-1} m_i = (\|m\|_1 - m \cdot \boldsymbol{\sigma})/2 > (\|m\|_1 - \|m\|_2^2 + N\varepsilon)/2 \ge N\varepsilon/2.$$

In both cases, the number of summands on the left-hand side is at least $N\varepsilon/2$ and, if we set

$$d := \frac{3d_1(\mu_m, \mu_{m'})}{\varepsilon},$$

then, by our assumption, $d < \varepsilon/2$. This means that the number of summands is greater than Nd. In the first case, the sum of the largest Nd values m_i corresponding to $\sigma_i = 1$ must be at least

$$\frac{Nd\varepsilon}{2} = \frac{3Nd_1(\mu_m, \mu_{m'})}{2}$$

and, therefore, if we flip the sign of σ_i corresponding to these largest values from $+$ to $-$, the value of $m \cdot \boldsymbol{\sigma} = \sum_{\sigma_i=1} m_i - \sum_{\sigma_i=-1} m_i$ will decrease by at least $3Nd_1(\mu_m, \mu_{m'})$. By the upper bound in (7.26), if we flip them consecutively, somewhere along the way we will have a configuration $\boldsymbol{\sigma}'$ such that $-N\varepsilon < \boldsymbol{\sigma}' \cdot m - m \cdot m < N\varepsilon$ (since $N\varepsilon > 1$, in one step we cannot jump from $N\varepsilon$ to $-N\varepsilon$). In other words, $\boldsymbol{\sigma}' \in B(m, \varepsilon)$ and, by construction, $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}'\|_1 \le 2dN$. The second

case is similar (we flip the sign of σ_i corresponding to the largest values from $-$ to $+$), and we showed that

$$\forall \boldsymbol{\sigma} \in B(m', \varepsilon),\ \ \exists \boldsymbol{\sigma}' \in B(m, \varepsilon) :\ \|\boldsymbol{\sigma}' - \boldsymbol{\sigma}\|_1 \le 2dN. \tag{7.28}$$

This implies that

$$\begin{aligned}&\Psi_N(m', \varepsilon, f, \zeta)\\ &\le \frac{1}{N}\mathbb{E}\log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\boldsymbol{\sigma}\in B(m,\varepsilon)} \sum_{\|\boldsymbol{\sigma}'-\boldsymbol{\sigma}\|_1\le 2dN} \exp \sum_{i\le N} (g_{f',i}(\alpha) + v(m'_i))\widetilde{\sigma}'_i,\end{aligned} \tag{7.29}$$

where we denote $\widetilde{\boldsymbol{\sigma}}' = \boldsymbol{\sigma}' - m'$. First of all, using (7.21) and the fact that $|\widetilde{\sigma}'_i| \le 2$, we can bound this by

$$\begin{aligned}&2\Delta_v(d_1(\mu_m, \mu_{m'}))\\ &\quad + \frac{1}{N}\mathbb{E}\log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\boldsymbol{\sigma}\in B(m,\varepsilon)} \sum_{\|\boldsymbol{\sigma}'-\boldsymbol{\sigma}\|_1\le 2dN} \exp \sum_{i\le N} (g_{f',i}(\alpha) + v(m_i))\widetilde{\sigma}'_i.\end{aligned}$$

The second term is equal to $\varphi(1)$ if we define

$$\begin{aligned}&\varphi(s) =\\ &\quad \frac{1}{N}\mathbb{E}\log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\boldsymbol{\sigma}\in B(m,\varepsilon)} \sum_{\|\boldsymbol{\sigma}'-\boldsymbol{\sigma}\|_1\le 2dN} \exp \sum_{i\le N} (g_{f',i}(\alpha) + v(m_i))(\widetilde{\sigma}_i + s(\widetilde{\sigma}'_i - \widetilde{\sigma}_i)).\end{aligned}$$

When $\|\boldsymbol{\sigma}' - \boldsymbol{\sigma}\|_1 \le 2dN$, we have

$$\|\widetilde{\boldsymbol{\sigma}}' - \widetilde{\boldsymbol{\sigma}}\|_1 \le \|\boldsymbol{\sigma}' - \boldsymbol{\sigma}\|_1 + \|m' - m\|_1 \le (d_1(\mu_m, \mu_{m'}) + 2d)N.$$

Therefore, differentiating $\varphi(s)$ and using Gaussian integration by parts, we get

$$\begin{aligned}|\varphi'(s)| &= \left|\frac{1}{N}\mathbb{E}\left\langle \sum_{i\le N} (g_{f',i}(\alpha) + v(m_i))(\widetilde{\sigma}'_i - \widetilde{\sigma}_i)\right\rangle_s\right|\\ &\le \left|\frac{1}{N}\mathbb{E}\left\langle \sum_{i\le N} g_{f',i}(\alpha)(\widetilde{\sigma}'_i - \widetilde{\sigma}_i)\right\rangle_s\right| + \|v\|_\infty(d_1(\mu_m, \mu_{m'}) + 2d)\\ &\le (c_f + \|v\|_\infty)(d_1(\mu_m, \mu_{m'}) + 2d),\end{aligned}$$

where $\langle\cdot\rangle_s$ denotes the Gibbs average along this interpolation. Lastly, if $\mathcal{I}(t)$ is the rate function of a Rademacher random variable and $\boldsymbol{\sigma}_0$ is any fixed vector in Σ_N, then

$$\begin{aligned}|\Psi_N(m, \varepsilon, \zeta) - \varphi(0)| &\le \frac{1}{N}\log \operatorname{card}\{\boldsymbol{\sigma} :\ \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_0\|_1 \le 2dN\}\\ &\le \log 2 - \mathcal{I}(1 - 2d) \le d\log\frac{1}{d}.\end{aligned}$$

Combining the above, we bounded $\Psi_N(m', \varepsilon, f, \zeta)$ by

$$\begin{aligned}&\Psi_N(m, \varepsilon, f, \zeta) + 2\Delta_v(d_1(\mu_m, \mu_{m'})) + \big(c_f + \|v\|_\infty\big)\big(d_1(\mu_m, \mu_{m'}) + 2d\big)\\ &\quad + d\log\frac{1}{d} = \Psi_N(m, \varepsilon, f, \zeta) + \Delta_f(d_1(\mu_m, \mu_{m'}), \varepsilon).\end{aligned}$$

Since the same inequality holds with m and m' interchanged, this proves (7.22).

The proof of (7.23) is similar. Using (7.28), we can write

$$\begin{aligned}&\mathbb{E}\,F^h_{N,1}(m',\varepsilon)\\ &\le \frac{1}{N}\mathbb{E}\log \sum_{\boldsymbol{\sigma}\in B(m,\varepsilon)} \sum_{\|\boldsymbol{\sigma}'-\boldsymbol{\sigma}\|_1\le 2dN} \exp\Big(H_N^{m'}(\widetilde{\boldsymbol{\sigma}}') + \sum_{i=1}^N v(m_i')\widetilde{\sigma}_i'\Big).\end{aligned} \tag{7.30}$$

We can handle the external field term as above and bound this by

$$\begin{aligned}&2\Delta_v(d_1(\mu_m,\mu_{m'})) + \|v\|_\infty(d_1(\mu_m,\mu_{m'})+2d)\\ &+\frac{1}{N}\mathbb{E}\log \sum_{\boldsymbol{\sigma}\in B(m,\varepsilon)} \sum_{\|\boldsymbol{\sigma}'-\boldsymbol{\sigma}\|_1\le 2dN} \exp\Big(H_N^{m'}(\widetilde{\boldsymbol{\sigma}}') + \sum_{i=1}^N v(m_i)\widetilde{\sigma}_i\Big).\end{aligned}$$

Next, we will replace $H_N^{m'}(\widetilde{\boldsymbol{\sigma}}')$ by $H_N^{m}(\widetilde{\boldsymbol{\sigma}})$ by using the interpolation

$$H_s(\widetilde{\boldsymbol{\sigma}}',\widetilde{\boldsymbol{\sigma}}) := \sqrt{s}H_N^{m'}(\widetilde{\boldsymbol{\sigma}}') + \sqrt{1-s}H_N^{m}(\widetilde{\boldsymbol{\sigma}}).$$

By Gaussian integration by parts, the error of this interpolation will be controlled by (twice) the maximum of the covariance

$$\frac{1}{N}\mathbb{E}\frac{\partial H_s(\widetilde{\boldsymbol{\sigma}}',\widetilde{\boldsymbol{\sigma}})}{\partial s}H_s(\widetilde{\boldsymbol{\rho}}',\widetilde{\boldsymbol{\rho}}) = \frac{1}{2}\big(\xi_a(\widetilde{R}'_{1,2}) - \xi_b(\widetilde{R}_{1,2})\big),$$

where

$$\begin{gathered}a = \frac{1}{N}\|m'\|_2^2, \quad b = \frac{1}{N}\|m\|_2^2, \quad \boldsymbol{\sigma},\boldsymbol{\rho}\in B(m,\varepsilon),\\ \|\boldsymbol{\sigma}'-\boldsymbol{\sigma}\|_1\le 2dN, \quad \|\boldsymbol{\rho}'-\boldsymbol{\rho}\|_1\le 2dN,\end{gathered}$$

and

$$\widetilde{R}'_{1,2} = \frac{1}{N}(\boldsymbol{\sigma}'-m')\cdot(\boldsymbol{\rho}'-m'), \quad \widetilde{R}_{1,2} = \frac{1}{N}(\boldsymbol{\sigma}-m)\cdot(\boldsymbol{\rho}-m).$$

Let us rewrite this as

$$\frac{1}{2}\big(\xi_a(\widetilde{R}'_{1,2}) - \xi_b(\widetilde{R}_{1,2})\big) = \frac{1}{2}\big(\xi_a(\widetilde{R}'_{1,2}) - \xi_a(\widetilde{R}_{1,2})\big) + \frac{1}{2}\big(\xi_a(\widetilde{R}_{1,2}) - \xi_b(\widetilde{R}_{1,2})\big),$$

and recall that $\xi_q(x) = \xi(x+q)-\xi(q)-\xi'(q)x$. The second term can be bounded by

$$\frac{1}{2}\big|\xi_a(\widetilde{R}_{1,2}) - \xi_b(\widetilde{R}_{1,2})\big| \le \frac{c_\xi}{2}|a-b| \le c_\xi d_1(\mu_m,\mu_{m'}).$$

To bound the first term, by the triangle inequality,

$$\begin{aligned}|\widetilde{R}'_{1,2} - \widetilde{R}_{1,2}| &\le \frac{2}{N}\big(\|\boldsymbol{\sigma}'-\boldsymbol{\sigma}\|_1 + \|\boldsymbol{\rho}'-\boldsymbol{\rho}\|_1 + 2\|m'-m\|_1\big)\\ &\le 4(d_1(\mu_m,\mu_{m'})+2d).\end{aligned}$$

Therefore, the first term can be bounded by $c_\xi(d_1(\mu_m,\mu_{m'})+2d)$. The rest of the argument is identical, so the proof of (7.23) is complete. □

We prove the following lemma by adapting an idea from Lemma 4 of [50].

LEMMA 7.6. *For any $\mu \in \mathcal{M}_{0,1}$ and $m = m^N$ satisfying $\mu_{m^N} \to \mu$, the limit*

$$\Psi(\mu, \varepsilon, f, \zeta) := \lim_{N\to\infty} \Psi_N(m^N, \varepsilon, f, \zeta) \tag{7.31}$$

exists and does not depend on the choice of the sequence (m^N).

We will denote

$$\Psi(\mu, f, \zeta) := \lim_{\varepsilon\downarrow 0} \Psi(\mu, \varepsilon, f, \zeta). \tag{7.32}$$

Consistently with (7.7), for $\zeta \in \mathcal{M}^d_{0,1-q}$ and $f = \xi_q$, we will denote

$$\Psi(m, \varepsilon, \zeta) := \Psi(m, \varepsilon, \xi_q, \zeta),\ \Psi(\mu, \zeta) := \Psi(\mu, \xi_q, \zeta). \tag{7.33}$$

PROOF. First, assume that μ is an atomic measure with rational weights. Suppose that for some $K, m_1, \dots, m_K$ is a sequence such that $\mu = \frac{1}{K}\sum_{i\le K} \delta_{m_i}$. For $i > K$, define m_i periodically $m_i = m_{i-K}$ and, for any N, define $m = m^N = (m_1, \dots, m_N)$, so that, clearly, $\mu_m \to \mu$.

Let $N_1, N_2 \ge 1$ be multiples of K and set $N = N_1 + N_2$. Note that

$$B(m^{N_1}, \varepsilon) \times B(m^{N_2}, \varepsilon) \subseteq B(m^N, \varepsilon). \tag{7.34}$$

Combined with standard properties of Ruelle probability cascades (see pp. 51–52 of [44]), this implies that

$$N\Psi_N(m^N, \varepsilon, f, \zeta) \tag{7.35}$$

is a superadditive function of N on multiples of K. One can easily verify from this that the limit $\lim_{N\to\infty} \Psi_N(m^N, \varepsilon, f, \zeta)$ exists.

Since the set of atomic measures with rational weights is dense in $\mathcal{M}_{0,1}$, by Lemma 7.5, the same limit exists for general μ, and it does not depend on the choice of m. Since (7.35) is decreasing in ε, the limit in (7.32) is well-defined. □

We will combine this result with Lemma 7.3 to obtain the following.

LEMMA 7.7. *If $\zeta \in \mathcal{M}^d_{0,1-q}$ and $\zeta' \in \mathcal{M}^d_{0,1-q'}$ are such that $\int_0^1 |\zeta(s) - \zeta'(s)|ds \le |q - q'|$, then, for any f_1, f_2 and $\mu, \mu' \in \mathcal{M}_{0,1}$,*

$$\begin{aligned} &|\Psi(\mu, \varepsilon, f_1, \zeta) - \Psi(\mu', \varepsilon, f_2, \zeta')| \\ &\quad \le \Delta_{f_1}\big(d_1(\mu, \mu'), \varepsilon\big) + 6f_1''(1)|q - q'| + \|f_1' - f_2'\|_\infty. \end{aligned} \tag{7.36}$$

PROOF. Combining Lemma 7.3 with the equation (7.22), we get

$$\begin{aligned} &\big|\Psi_N(m, \varepsilon, f_1, \zeta) - \Psi_N(m', \varepsilon, f_2, \zeta')\big| \\ &\quad \le \Delta_{f_1}\big(d_1(\mu_m, \mu_{m'}), \varepsilon\big) + 6f_1''(1)|q - q'| + \|f_1' - f_2'\|_\infty. \end{aligned}$$

Letting $\mu_m \to \mu$ and $\mu_{m'} \to \mu'$ and using the previous lemma finishes the proof. □

7.2 Proof of the upper bound of Proposition 7.1

Fix some $\mu \in \mathcal{M}_{0,1}$ and $v \in V$ throughout this subsection and let $m = m^N$ be a sequence with $\mu_{m^N} \to \mu$. Take $q = \int x^2 \, d\mu(x)$. In view of Lemma 7.6, it will be enough to show that for arbitrary $\zeta \in \mathcal{M}^d_{0,1-q}$ and $\varepsilon_N \to 0$,

$$\limsup_{N\to\infty} \mathbb{E} F^h_{N,1}(m^N, \varepsilon_N) \le \limsup_{N\to\infty} \Psi_N(m^N, \varepsilon_N, \zeta) - \Upsilon(\zeta). \tag{7.37}$$

Once we proved this, we can simply choose $\varepsilon_N \to 0$ slowly enough so that

$$\limsup_{N\to\infty} \mathbb{E} F^h_{N,1}(m^N, \varepsilon_N) = \lim_{\varepsilon\downarrow 0} \limsup_{N\to\infty} \mathbb{E} F^h_{N,1}(m^N, \varepsilon)$$

and, simultaneously,

$$\limsup_{N\to\infty} \Psi_N(m^N, \varepsilon_N, \zeta) = \lim_{\varepsilon\downarrow 0} \lim_{N\to\infty} \Psi_N(m^N, \varepsilon, \zeta),$$

which is equal to $\Psi(\mu, \zeta)$, by Lemma 7.6. This will finish the proof of the upper bound of Proposition 7.1.

Our proof of (7.37) uses Guerra's interpolation method and is almost identical to the proof of that for mixed p-spin models on Σ_N as in sections 3.2–3.4 in [44] except that we need to work with the shifted coordinates $\widetilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - m$ instead of $\boldsymbol{\sigma}$ and replace the cube Σ_N by the band $B(m, \varepsilon_N)$. Similar to [44, eq. (3.45)], we define, for $t \in [0, 1]$, the interpolating Hamiltonian

$$H_{N,t}(\widetilde{\boldsymbol{\sigma}}, \alpha) = \sqrt{t} H^m_N(\widetilde{\boldsymbol{\sigma}}) + \sqrt{1-t} \sum_{i=1}^N g_{\xi'_q,i}(\alpha)\widetilde{\sigma}_i + \sqrt{t} \sum_{i=1}^N g_{\theta_q,i}(\alpha) + \sum_{i=1}^N v(m_i)\widetilde{\sigma}_i,$$

indexed by $(\widetilde{\boldsymbol{\sigma}}, \alpha) \in (B(m, \varepsilon_N) - m) \times \mathbb{N}^r$, where the Gaussian processes $g_{\xi'_q,i}$ and $g_{\theta_q,i}$ are i.i.d. copies of the processes defined in (7.4). Exactly as in [44, eq. (3.18)], set

$$g(\widetilde{\boldsymbol{\sigma}}) = \sum_{p\ge 1} 2^{-p} x_p g_p(\widetilde{\boldsymbol{\sigma}}), \quad g_p(\widetilde{\boldsymbol{\sigma}}) = \frac{1}{N^{p/2}} \sum_{i_1,\dots,i_p=1}^N g'_{i_1,\dots,i_p} \widetilde{\sigma}_{i_1} \cdots \widetilde{\sigma}_{i_p}, \tag{7.38}$$

$g'_{i_1,\dots,i_p}$ are i.i.d. standard Gaussian variables, and $x = (x_p)_{p\ge 1}$ is a sequence of i.i.d. uniform random variables on $[1, 2]$. Let s_N be a sequence such that $s_N \to \infty$ and $s_N^2/N \to 0$. For $t \in [0, 1]$, define the interpolating free energy by

$$\varphi_N(t) = \frac{1}{N} \mathbb{E}_x \mathbb{E} \log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\boldsymbol{\sigma}\in B(m,\varepsilon_N)} \exp\big(H_{N,t}(\widetilde{\boldsymbol{\sigma}}, \alpha) + s_N g(\widetilde{\boldsymbol{\sigma}})\big). \tag{7.39}$$

Here and hereinafter, $\mathbb{E}_x$ means the expectation with respect to the randomness x only. Denote by $G_{N,t}(\widetilde{\boldsymbol{\sigma}}, \alpha)$ the Gibbs measure and by $\langle\cdot\rangle_t$ the Gibbs average

associated to this free energy. Observe that since $s_N^2/N \to 0$, the term $s_N g(\tilde{\boldsymbol{\sigma}})$ plays the role as a vanishing perturbation such that

$$\begin{aligned} \varphi_N(1) &= \mathbb{E} F_{N,1}^h(m) + \Upsilon(\zeta) + o(1), \\ \varphi_N(0) &= \Psi_N(m, \varepsilon_N, \zeta) + o(1). \end{aligned} \tag{7.40}$$

In order to compare these two sides, an application of the Gaussian integration by parts (see [44, theorem 3.5]) implies that, as $N \to \infty$,

$$\varphi_N'(t) = -\frac{1}{2}\mathbb{E}_x\mathbb{E}\Big\langle\Big(\xi_q(R(\tilde{\boldsymbol{\sigma}}^1, \tilde{\boldsymbol{\sigma}}^2)) - R(\tilde{\boldsymbol{\sigma}}^1, \tilde{\boldsymbol{\sigma}}^2)\xi_q'(q_{\alpha_1\wedge\alpha_2}) + \theta_q(q_{\alpha_1\wedge\alpha_2})\Big)\Big\rangle_t + o(1).$$

The term $o(1)$ comes from the bound on the expression involving self-overlaps,

$$\xi_q(R(\tilde{\boldsymbol{\sigma}}^1, \tilde{\boldsymbol{\sigma}}^1)) - R(\tilde{\boldsymbol{\sigma}}^1, \tilde{\boldsymbol{\sigma}}^1)\xi_q'(q_{\alpha_1\wedge\alpha_1}) + \theta_q(q_{\alpha_1\wedge\alpha_1}).$$

Indeed, the fact that $\varepsilon_N \to 0$ and $\mu_{m^N} \to \mu$ ensures, by (7.19), that $R(\tilde{\boldsymbol{\sigma}}^1, \tilde{\boldsymbol{\sigma}}^1) \approx 1-q$, and by our choice of $q_r = 1-q$, we have $q_{\alpha_1\wedge\alpha_1} = 1-q$. Because the self-overlap is nearly constant, the proof of the extended Ghirlanda-Guerra identities in the average sense, as well as Talagrand's positivity principle, in [44] requires no modifications and theorem 3.4 in [44] implies that

$$\lim_{N\to\infty} \mathbb{E}_x\mathbb{E}\langle \mathbf{1}(R(\tilde{\boldsymbol{\sigma}}^1, \tilde{\boldsymbol{\sigma}}^2) \le -\varepsilon)\rangle_t = 0 \ \ \forall \varepsilon > 0.$$

This together with the fact that ξ_q is a convex function on $\mathbb{R}_+$ implies that

$$\limsup_{N\to\infty} \varphi_N'(t) \le 0.$$

Consequently, the asserted inequality follows from (7.40). □

7.3 Proof of the lower bound of Proposition 7.1

Let us consider an atomic measure μ with finitely many jumps and rational weights, and let

$$q = \int x^2\, d\mu(x), \gamma = \int x(1-x) d\mu(x).$$

Denote

$$\delta(\varepsilon, \gamma) := (c_\xi + \|v\|_\infty)\Big(\mathbf{1}(\gamma > 0)\frac{\varepsilon}{\gamma}\log\frac{\gamma}{\varepsilon} + \mathbf{1}(\gamma = 0)\varepsilon\log\frac{1}{\varepsilon}\Big), \tag{7.41}$$

for some constant c_ξ that will be determined in the proof below. Let $\varepsilon > 0$ be such that $\varepsilon < \gamma^2/2$, when $\gamma > 0$. We will show that if $\mu_m \to \mu$, then

$$\liminf_{N\to\infty} \mathbb{E} F_{N,1}^h(m, \varepsilon) \ge \inf_{\zeta\in\mathcal{M}_{0,1-q}^d} \big(\Psi(\mu, \varepsilon, \zeta) - \Upsilon(\zeta)\big) - \varepsilon - \delta(\varepsilon, \gamma). \tag{7.42}$$

Before we prove this, let us show why this implies the lower bound in Proposition 7.1.

LEMMA 7.8. *If* (7.42) *holds for atomic measures* μ *with rational weights, then the lower bound in Proposition* 7.1 *holds for all* $\mu \in \mathcal{M}_{0,1}$.

PROOF. Take $\mu \in \mathcal{M}_{0,1}$ with $\gamma = \int x(1-x)d\mu(x)$ and $q = \int x^2\, d\mu(x)$. We can find an atomic $\mu' \in \mathcal{M}_{0,1}$ with rational weights and $d_1(\mu, \mu')$ as small as we wish, but we can also make sure that if $\gamma = 0$, then $\gamma' = \int x(1-x)d\mu'(x) = 0$, and if $\gamma > 0$, then $\gamma' > \gamma/2$. Let $q' = \int x^2\, d\mu'(x)$. When $d_1(\mu, \mu') < \varepsilon^2/6$, (7.23) implies that (assuming (7.42))

$$\begin{aligned}
&\liminf_{N\to\infty} \mathbb{E} F^h_{N,1}(m, \varepsilon) \\
&\quad \geq \liminf_{N\to\infty} \mathbb{E} F^h_{N,1}(m', \varepsilon) - \Delta_\xi\big(d_1(\mu, \mu'), \varepsilon\big) \\
&\quad \geq \inf_{\zeta' \in \mathcal{M}^d_{0,1-q'}} \big(\Psi(\mu', \varepsilon, \zeta') - \Upsilon(\zeta')\big) - \Delta_\xi\big(d_1(\mu, \mu'), \varepsilon\big) - \varepsilon - \delta(\varepsilon, \gamma).
\end{aligned}$$

For any $\zeta' \in \mathcal{M}^d_{0,1-q'}$ we can find $\zeta \in \mathcal{M}^d_{0,1-q}$ such that $\int_0^1 |\zeta(s) - \zeta'(s)|ds \leq |q - q'|$. Then by Lemma 7.2,

$$|\Upsilon(\zeta) - \Upsilon(\zeta')| = |\Upsilon(\xi_q, \zeta) - \Upsilon(\xi_{q'}, \zeta')| \leq c_\xi |q - q'| \tag{7.43}$$

, and, using Lemma 7.7 with $f_1 = \xi_q$ and $f_2 = \xi_{q'}$, we can bound the above lower limit from below by

$$\inf_{\zeta \in \mathcal{M}^d_{0,1-q}} \big(\Psi(\mu, \varepsilon, \zeta) - \Upsilon(\zeta)\big) - 2\Delta_{\xi_q}(d_1(\mu, \mu'), \varepsilon) - \varepsilon - \delta(\varepsilon, \gamma) - c_\xi |q - q'|.$$

Here we also used that $\|\xi'_q - \xi'_{q'}\|_\infty \leq c_\xi |q - q'|$. Letting $\mu' \to \mu$ (so that $q' \to q$) and then letting $\varepsilon \downarrow 0$ finishes the proof. □

We now proceed with the proof of (7.42). Suppose that for some $K, m_1, \ldots, m_K$ is a sequence such that $\mu = K^{-1} \sum_{i \leq K} \delta_{m_i}$. For $i > K$, define m_i periodically $m_i = m_{i-K}$ and define $m = m^N = (m_1, \ldots, m_N)$, so that $\mu_m \to \mu$. Since

$$\frac{\|m^N\|^2}{N} \to \int x^2\, d\mu(x) = q,$$

we have

$$\begin{aligned}
\mathbb{E} F^h_{N,1}(m, \varepsilon) &= \frac{1}{N} \mathbb{E} \log \sum_{\boldsymbol{\sigma} \in B(m^N, \varepsilon)} \exp\left(H^m_N(\widetilde{\boldsymbol{\sigma}}) + \sum_{i \leq N+M} v(m_i)\widetilde{\sigma}_i \right) \\
&\approx \frac{1}{N} \mathbb{E} \log \sum_{\boldsymbol{\sigma} \in B(m^N, \varepsilon)} \exp\left(H^q_N(\widetilde{\boldsymbol{\sigma}}) + \sum_{i \leq N+M} v(m_i)\widetilde{\sigma}_i \right)
\end{aligned}$$

where $H^q_N(\widetilde{\boldsymbol{\sigma}})$ is the Hamiltonian with the covariance $N\xi_q(R(\widetilde{\boldsymbol{\sigma}}^1, \widetilde{\boldsymbol{\sigma}}^2))$ with the function ξ_q defined as in (3.5). Recall that $H^m_N(\widetilde{\boldsymbol{\sigma}})$ was defined exactly as $H^q_N(\widetilde{\boldsymbol{\sigma}})$ only with q given by $\|m^N\|^2/N$, and the standard interpolation argument (as in the proof of Lemma 7.5) shows that we can replace $\|m^N\|^2/N$ by its limit q to make sure we are working with the 'same' Hamiltonian along the entire sequence.

The proof of the lower bound will be based on a standard cavity computation. Fix $\varepsilon > 0$ and some integer $M \geq 1$ and write

$$\begin{aligned}
&\liminf_{N\to\infty} \mathbb{E} F^h_{N,1}(m,\varepsilon) \\
&\quad \geq \liminf_{N\to\infty} \frac{1}{M}\Bigg(\mathbb{E}\log \sum_{\sigma\in B(m^{N+M},\varepsilon)} \exp\Bigg(H^q_{N+M}(\tilde{\sigma}) + \sum_{i\leq N+M} v(m_i)\tilde{\sigma}_i \Bigg) \\
&\qquad\qquad - \mathbb{E}\log \sum_{\sigma\in B(m^N,\varepsilon)} \exp\Bigg(H^q_N(\tilde{\sigma}) + \sum_{i\leq N} v(m_i)\tilde{\sigma}_i \Bigg)\Bigg).
\end{aligned} \tag{7.44}$$

Notice that, since the lower limit on the left-hand side does not change if we take it over N proportional to K, we can take the lower limit on the right-hand side also over such N. In particular, by periodicity,

$$m_{N+i} = m_i \quad \text{and} \quad \mu_{m^N} = \mu. \tag{7.45}$$

By an abuse of notation, let $m^M = (m^M_i)_{i\leq M}$ denote the vector with elements $m^M_i = m_{N+i} = m_i$. The latter are the so-called cavity coordinates. Note that

$$B(m^N,\varepsilon)\times B(m^M,\varepsilon) \subseteq B(m^{N+M},\varepsilon).$$

Assume henceforth that M is large enough so that all the bands above are nonempty for large N. Denoting by $H^q_{N+M}(\tilde{\sigma},\tilde{\rho})$ the value of the Hamiltonian H^q_{N+M} at the vector obtained by concatenation of $\tilde{\sigma} = \sigma - m^N$ and $\tilde{\rho} = \rho - m^M$, (7.44) is bounded from below by

$$\begin{aligned}
&\liminf_{N\to\infty} \frac{1}{M}\Bigg(\mathbb{E}\log \sum_{\substack{\sigma\in B(m^N,\varepsilon)\\ \rho\in B(m^M,\varepsilon)}} \exp\Bigg(H^q_{N+M}(\tilde{\sigma},\tilde{\rho}) + \sum_{i\leq N} v(m_i)\tilde{\sigma}_i + \sum_{i\leq M} v(m_i)\tilde{\rho}_i \Bigg) \\
&\qquad - \mathbb{E}\log \sum_{\sigma\in B(m^N,\varepsilon)} \exp\Bigg(H^q_N(\tilde{\sigma}) + \sum_{i\leq N} v(m_i)\tilde{\sigma}_i \Bigg)\Bigg).sBy
\end{aligned} \tag{7.46}$$

By replacing $B(m^N,\varepsilon)$ by $B(m^N,\varepsilon_N)$ in the first sum, with some sequence $\varepsilon_N \to 0$, we only reduce (7.46). To replace $B(m^N,\varepsilon)$ by $B(m^N,\varepsilon_N)$ in the second term, we can argue as in Lemma 7.5, as follows. If $\sigma \in B(m,\varepsilon)\setminus B(m,\varepsilon_N)$, then there are two possibilities:

$$N\varepsilon_N < \sigma\cdot m - m\cdot m < N\varepsilon, \tag{7.47}$$

$$N\varepsilon_N < m\cdot m - \sigma\cdot m < N\varepsilon. \tag{7.48}$$

Since, by (7.45), $\mu_m = \mu$ and, therefore,

$$\frac{1}{N}\big(\|m\|_1 - \|m\|_2^2\big) = \int x(1-x)d\mu_m(x) = \gamma.$$

We will argue differently in the case when $\gamma > 0$ or $\gamma = 0$.

If $\gamma > 0$ (arguing as below (7.26), (7.27) above) in the first case,

$$\sum_{\sigma_i=1} m_i = (\|m\|_1 + m\cdot\boldsymbol{\sigma})/2 > \big(\|m\|_1 + \|m\|_2^2 + N\varepsilon_N\big)/2 \geq N\gamma/2 \tag{7.49}$$

and, in the second case,

$$\sum_{\sigma_i=-1} m_i = (\|m\|_1 - m\cdot\boldsymbol{\sigma})/2 > \big(\|m\|_1 - \|m\|_2^2 + N\varepsilon_N\big)/2 \geq N\gamma/2. \tag{7.50}$$

If $N\varepsilon_N > 1$, $\varepsilon < \gamma^2/2$ and $d := \varepsilon/\gamma$, then we can argue exactly as in Lemma 7.5 that

$$\forall \boldsymbol{\sigma} \in B(m,\varepsilon),\ \ \exists \boldsymbol{\sigma}' \in B(m,\varepsilon_N) :\ \|\boldsymbol{\sigma}' - \boldsymbol{\sigma}\|_1 \leq 2dN \tag{7.51}$$

and, therefore,

$$\begin{aligned}
&\mathbb{E}\log \sum_{\boldsymbol{\sigma}\in B(m^N,\varepsilon)} \exp\Big(H_N^q(\widetilde{\boldsymbol{\sigma}}) + \sum_{i\leq N} v(m_i)\widetilde{\sigma}_i\Big) \\
&\quad\leq \mathbb{E}\log \sum_{\boldsymbol{\sigma}\in B(m^N,\varepsilon_N)} \exp\Big(H_N^q(\widetilde{\boldsymbol{\sigma}}) + \sum_{i\leq N} v(m_i)\widetilde{\sigma}_i\Big) + (c_\xi + \|v\|_\infty)\frac{\varepsilon}{\gamma}\log\frac{\gamma}{\varepsilon}.
\end{aligned}$$

Using this, we can bound (7.46) from below by

$$\begin{aligned}
\liminf_{N\to\infty}\frac{1}{M}\Bigg(&\mathbb{E}\log \sum_{\substack{\boldsymbol{\sigma}\in B(m^N,\varepsilon_N)\\ \boldsymbol{\rho}\in B(m^M,\varepsilon)}} \exp\Big(H_{N+M}^q(\widetilde{\boldsymbol{\sigma}},\widetilde{\boldsymbol{\rho}}) + \sum_{i\leq N} v(m_i)\widetilde{\sigma}_i + \sum_{i\leq M} v(m_i)\widetilde{\rho}_i\Big) \\
&- \mathbb{E}\log \sum_{\boldsymbol{\sigma}\in B(m^N,\varepsilon_N)} \exp\Big(H_N^q(\widetilde{\boldsymbol{\sigma}}) + \sum_{i\leq N} v(m_i)\widetilde{\sigma}_i\Big)\Bigg) \\
&- (c_\xi + \|v\|_\infty)\frac{\varepsilon}{\gamma}\log\frac{\gamma}{\varepsilon}.
\end{aligned} \tag{7.52}$$

Now, let us consider the case $\gamma = 0$, when all m_i are equal to 0 or 1. Since $\boldsymbol{\sigma}\cdot m - m\cdot m \leq N(\|m\|_1 - \|m\|_2^2) = 0$, the first case (7.47) is not possible. The second case (7.48) can be rewritten as

$$N\varepsilon_N < m\cdot m - \boldsymbol{\sigma}\cdot m = 2\sum_{\sigma_i=-1} m_i = 2\sum_{\sigma_i=-1}\mathbf{1}(m_i = 1) < N\varepsilon. \tag{7.53}$$

First of all, if we flip all $\sigma_i = -1$ corresponding to $m_i = 1$ to $+1$, then $m\cdot m - \boldsymbol{\sigma}\cdot m$ will become 0. If we flip them consecutively, then somewhere along the way we will find $\boldsymbol{\sigma}' \in B(m,\varepsilon_N)$. On the other hand, the second inequality in (7.53) implies that the number of such $m_i = 1$ is bounded by $N\varepsilon/2$. Therefore, we need to flip at most $N\varepsilon/2$ coordinates, which proves that

$$\forall \boldsymbol{\sigma} \in B(m,\varepsilon),\ \ \exists \boldsymbol{\sigma}' \in B(m,\varepsilon_N) :\ \|\boldsymbol{\sigma}' - \boldsymbol{\sigma}\|_1 \leq \varepsilon N/2. \tag{7.54}$$

Therefore,

$$\mathbb{E}\log \sum_{\sigma\in B(m^N,\varepsilon)} \exp\Big(H_N^q(\tilde\sigma)+\sum_{i\le N} v(m_i)\tilde\sigma_i\Big)$$
$$\le \mathbb{E}\log \sum_{\sigma\in B(m^N,\varepsilon_N)} \exp\big(H_N^q(\tilde\sigma)+\sum_{i\le N} v(m_i)\tilde\sigma_i\big) + (c_\xi+\|v\|_\infty)\varepsilon\log\frac{1}{\varepsilon}$$

and (7.46) is bounded from below by

$$\liminf_{N\to\infty}\frac{1}{M}\Big(\mathbb{E}\log \sum_{\substack{\sigma\in B(m^N,\varepsilon_N)\\ \rho\in B(m^M,\varepsilon)}} \exp\big(H_{N+M}^q(\tilde\sigma,\tilde\rho)+\sum_{i\le N} v(m_i)\tilde\sigma_i+\sum_{i\le M} v(m_i)\tilde\rho_i\big) \tag{7.55}$$
$$-\mathbb{E}\log \sum_{\sigma\in B(m^N,\varepsilon_N)} \exp\big(H_N^q(\tilde\sigma)+\sum_{i\le N} v(m_i)\tilde\sigma_i\big)\Big) - (c_\xi+\|v\|_\infty)\varepsilon\log\frac{1}{\varepsilon}.$$

Recalling the notation (7.41)), both cases can be combined as

$$\liminf_{N\to\infty}\frac{1}{M}\Big(\mathbb{E}\log \sum_{\substack{\sigma\in B(m^N,\varepsilon_N)\\ \rho\in B(m^M,\varepsilon)}} \exp\big(H_{N+M}^q(\tilde\sigma,\tilde\rho)+\sum_{i\le N} v(m_i)\tilde\sigma_i+\sum_{i\le M} v(m_i)\tilde\rho_i\big) \tag{7.56}$$
$$-\mathbb{E}\log \sum_{\sigma\in B(m^N,\varepsilon_N)} \exp\big(H_N^q(\tilde\sigma)+\sum_{i\le N} v(m_i)\tilde\sigma_i\big)\Big) - \delta(\varepsilon,\gamma).$$

The advantage of working with (7.56) instead of (7.46) is that the self-overlap $R(\tilde\sigma,\tilde\sigma)$ converges uniformly over $\sigma\in B(m,\varepsilon_N)$ to $1-q$, a fact that will later be important when we invoke the Ghirlanda-Guerra identities.

Next we use the standard cavity computation, known as the Aizenman-Sims-Starr scheme [1] (see, e.g., [44] or [14]). Consider the Hamiltonian

$$H_{N,M}^q(\tilde\sigma) := \sum_{p\ge 2}\frac{\beta_p(q)}{(N+M)^{(p-1)/2}} \sum_{1\le i_1,\dots,i_p\le N} g_{i_1,\dots,i_p}\tilde\sigma_{i_1}\cdots\tilde\sigma_{i_p}, \tag{7.57}$$

with the covariance

$$\mathbb{E}H_{N,M}^q(\tilde\sigma^1)H_{N,M}^q(\tilde\sigma^2) = (N+M)\xi_q\Big(\frac{N}{N+M}R(\tilde\sigma^1,\tilde\sigma^2)\Big)$$

and independent Hamiltonians $z(\tilde\sigma)$ and $y(\tilde\sigma)$ with covariances

$$\mathbb{E}z(\tilde\sigma^1)z(\tilde\sigma^2) = \xi_q'\big(R(\tilde\sigma^1,\tilde\sigma^2)\big),$$
$$\mathbb{E}y(\tilde\sigma^1)y(\tilde\sigma^2) = \theta_q'\big(R(\tilde\sigma^1,\tilde\sigma^2)\big),$$

where $\theta_q(x) = x\xi_q'(x)-\xi_q(x)$. Let $z_i(\tilde\sigma)$ be independent copies of $z(\tilde\sigma)$ for $i\ge 1$. We denote by $G_{M,N}(\tilde\sigma)$ the Gibbs measure proportional to

$$\mathbf{1}_{B(m^N,\varepsilon_N)}(\sigma)\exp\Big(H_{N,M}^q(\tilde\sigma)+\sum_{i\le N} v(m_i)\tilde\sigma_i\Big) \tag{7.58}$$

and by $\langle\cdot\rangle_{N,M}$ its average. If we denote

$$T_{N,M} := \frac{1}{M}\mathbb{E}\log\Big\langle \sum_{\rho\in B(m^M,\varepsilon)} \exp\sum_{i\le M}\big(\widetilde{\rho}_i z_i(\widetilde{\sigma}) + v(m_i)\widetilde{\rho}_i\big)\Big\rangle_{N,M} - \frac{1}{M}\mathbb{E}\log\big\langle \exp\sqrt{M}\, y(\widetilde{\sigma})\big\rangle_{N,M}, \tag{7.59}$$

then a straightforward interpolation argument can be used to rewrite (7.56) as

$$\liminf_{N\to\infty} T_{N,M} - \delta(\varepsilon,\gamma). \tag{7.60}$$

Recall the perturbation $s_N g(\widetilde{\sigma})$ from (7.39). Everywhere above we could have replaced H_N^q by the perturbed Hamiltonian

$$H_{N,q}^{\mathrm{pert}}(\widetilde{\sigma}) = H_N^q(\widetilde{\sigma}) + s_N g(\widetilde{\sigma}), \tag{7.61}$$

with $s_N = N^{1/3}$ (here we can take any power strictly between $1/4$ and $1/2$). Then 1 can still show (see section 3.5 in [44]) that (7.60) still holds uniformly over the choice of (x_p) in the perturbation term $g(\widetilde{\sigma})$, with the Gibbs measure (7.58) modified by

$$G_{N,M}^{\mathrm{pert}}(\widetilde{\sigma}) \sim \mathbf{1}_{B(m^N,\varepsilon_N)}(\sigma)\exp\Big(H_{N,M}^q(\widetilde{\sigma}) + \sum_{i\le N} v(m_i)\widetilde{\sigma}_i + s_N g(\widetilde{\sigma})\Big).$$

Moreover, we can choose the parameters in the perturbation term $x_p^{N,M}$ in such a way that the above Gibbs measure satisfies the Ghirlanda-Guerra identities, asymptotically. This chain of arguments is standard, and we refer the reader to section 3.6 of [44] for more details.

Next, consider a subsequence along which the lower limit in (7.60) is obtained, and take a further subsequence along which the array of overlaps $(R(\widetilde{\sigma}^\ell\cdot\widetilde{\sigma}^{\ell'}))_{\ell,\ell'\ge 1}$ of configurations sampled from $G_{N,M}^{\mathrm{pert}}$ converges in the sense of finite-dimensional distributions. By the main result of [43], the Ghirlanda-Guerra identities imply that the limiting array of overlaps is ultrametric and it can be approximated (in the sense of finite-dimensional distributions) by the overlaps generated by a sequence of Ruelle probability cascades, say corresponding to a sequence of distribution functions ζ_n with finitely many atoms, as in (7.3).

Recall $\Psi_M(m,\varepsilon,\zeta)$ and $\Upsilon(\zeta)$ from (7.7) and (7.8). Denote

$$T_{\zeta,M} := \Psi_M(m^M,\varepsilon,\zeta) - \Upsilon(\zeta).$$

Recall the notation $T_{N,M}$ from (7.59). Note that the covariance function of the Gaussian processes $(g_{\xi_q',i}, g_{\theta_q})$ and (z_i, y) used in the definition of $T_{\zeta,M}$ and $T_{N,M}$ has the same dependence on the overlap as the variables α and $\widetilde{\sigma}$, respectively. Combining this with a straightforward generalization of [44, theorem 1.3], we have that there exists $\zeta^M \in \mathcal{M}_{0,1-q}^d$ such that

$$|\liminf_{N\to\infty} T_{N,M} - T_{\zeta^M,M}| < \varepsilon/2.$$

Finally, take a subsequence of (ζ^M) that converges to some limit distribution ζ'. By Lemma 7.2, if we choose $\zeta \in \mathcal{M}^d_{0,1-q}$ such that $d_1(\zeta', \zeta)$ is small enough, we can ensure that $|T_{\zeta^M,M} - T_{\zeta,M}| < \varepsilon/2$ for large enough M. This together with (7.60) gives that

$$\liminf_{N\to\infty} \mathbb{E} F^h_{N,1}(m,\varepsilon) \geq \Psi_M(m^M, \varepsilon, \zeta) - \Upsilon(\zeta) - \varepsilon - \delta(\varepsilon,\gamma),$$

for M large enough. Taking the limit $M \to \infty$ and using Lemma 7.6 proves (7.42). □

7.4 Proof of Theorem 3.2

Let us recall the definition of $\Psi_N(m,\varepsilon,\zeta)$ in (7.7),

$$\Psi_N(m,\varepsilon,\zeta) = \frac{1}{N}\mathbb{E}\log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\sigma\in B(m,\varepsilon)} \exp \sum_{i\leq N} (g_{\xi'_q,i}(\alpha) + v(m_i))\widetilde{\sigma}_i, \tag{7.62}$$

for $\zeta \in \mathcal{M}^d_{0,1-q}$. For $a \in [0,1]$ and $(t,x) \in [0,1-q]\times\mathbb{R}$, let $\Phi_{a,\zeta}(t,x)$ be the solution of

$$\partial_t \Phi_{a,\zeta} = -\frac{\xi''_q(t)}{2}\big(\partial_{xx}\Phi_{a,\zeta} + \zeta(t)(\partial_x \Phi_{a,\zeta})^2\big)$$

with the boundary condition

$$\Phi_{a,\zeta}(1-q,x) = \log \sum_{\sigma=\pm 1} e^{(\sigma-a)x} = \log 2 - ax + \log\cosh x. \tag{7.63}$$

By Proposition 7.1, in order to finish the proof of Theorem 3.2, we need to prove the following.

PROPOSITION 7.9. *For any $\zeta \in \mathcal{M}^d_{0,1-q}$, and $\mu \in \mathcal{M}_{0,1}$ with $\int x^2\, d\mu(x) = q$,*

$$\lim_{\varepsilon\downarrow 0}\lim_{N\to\infty} \Psi_N(m,\varepsilon,\zeta) = \inf_{\lambda\in\mathbb{R}} \int \Phi_{a,\zeta}(0, \lambda a + v(a))d\mu(a) \tag{7.64}$$

for any sequence $m = m^N$ such that $\mu_m \to \mu$.

The infimum in the Proposition 7.1 is taken over atomic $\zeta \in \mathcal{M}^d_{0,1-q}$, but, since the boundary condition satisfies

$$|\partial_x \Phi_{a,\zeta}(1-q,x)| \leq 2, |\partial_{xx}\Phi_{a,\zeta}(1-q,x)| \leq 1, \tag{7.65}$$

one can show (using the standard argument of Guerra mentioned in Lemma 7.2 above) that the functional $\zeta \to \Phi_{a,\zeta}(0,x)$ is uniformly d_1-Lipschitz over all parameters, and the infimum can be taken over all ζ.

Since for the rest of the section $\zeta \in \mathcal{M}^d_{0,1-q}$ is fixed, we will omit it and write (7.62) as

$$\Psi_N(m,\varepsilon) = \frac{1}{N}\mathbb{E}\log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\sigma\in B(m,\varepsilon)} \exp \sum_{i\leq N} (g_{\xi'_q,i}(\alpha) + v(m_i))\widetilde{\sigma}_i. \tag{7.66}$$

For $\lambda \in \mathbb{R}$ and $A \subseteq \Sigma_N$, let us define

$$\begin{aligned}\Psi_N(m;\lambda,A) &= \frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\sigma\in A}\exp\Big[\sum_{i\le N}\big(g_{\xi'_q,i}(\alpha)+v(m_i)\big)\widetilde{\sigma}_i + \lambda N R(\widetilde{\sigma},m)\Big] \\ &= \frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\sigma\in A}\exp\sum_{i\le N}\big(g_{\xi'_q,i}(\alpha)+\lambda m_i+v(m_i)\big)\widetilde{\sigma}_i.\end{aligned} \tag{7.67}$$

For $\varepsilon > 0$, recall that $B(m,\varepsilon) = \{\boldsymbol{\sigma} : |R(\widetilde{\boldsymbol{\sigma}},m)| < \varepsilon\}$ and let

$$B^+(m,\varepsilon) = \{\boldsymbol{\sigma} : R(\widetilde{\boldsymbol{\sigma}},m) \ge \varepsilon\},\quad B^-(m,\varepsilon) = \{\boldsymbol{\sigma} : R(\widetilde{\boldsymbol{\sigma}},m) \le -\varepsilon\}.$$

Note that for $\gamma \ge 0$,

$$\Psi_N(m;\lambda, B^{\pm}(m,\varepsilon)) \le \Psi_N(m;\lambda\pm\gamma,\Sigma_N) - \gamma\varepsilon. \tag{7.68}$$

Our strategy to prove Proposition 7.9 will be to show that, with the choice of $\lambda = \lambda_0$ corresponding to the minimizer in (7.64), for any $\varepsilon > 0$, the quantities $\Psi_N(m;\lambda, B^{\pm}(m,\varepsilon))$ will be strictly smaller than $\Psi_N(m;\lambda,\Sigma_N)$, which will imply (via concentration) that $\Psi_N(m;\lambda_0, B(m,\varepsilon)) \approx \Psi_N(m;\lambda_0,\Sigma_N)$. To achieve this goal, we need two auxiliary lemmas.

LEMMA 7.10. *For any $\lambda > 0$, we have*

$$\Psi_N(m;\lambda,\Sigma_N) = \int \Phi_{a,\xi}(0,\lambda a + v(a))\,d\mu_m(a) \tag{7.69}$$

and

$$\big|\Psi_N(m,\varepsilon) - \Psi_N(m;\lambda,B(m,\varepsilon))\big| \le |\lambda|\varepsilon. \tag{7.70}$$

PROOF. Recall the notation in (7.63). Then, by the standard properties of the Ruelle probability cascades,

$$\begin{aligned}&\Psi_N(m;\lambda,\Sigma_N) \\ &\quad= \frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\sigma\in\Sigma_N}\exp\sum_{i\le N}\big(g_{\xi'_q,i}(\alpha)+\lambda m_i+v(m_i)\big)\widetilde{\sigma}_i \\ &\quad= \frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha \prod_{i\le N}\sum_{\sigma=\pm1}\exp(\sigma-m_i)\big(g_{\xi'_q,i}(\alpha)+\lambda m_i+v(m_i)\big) \\ &\quad= \frac{1}{N}\sum_{i\le N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha \exp\Phi_{m_i,\xi}\big(1-q, g_{\xi'_q,i}(\alpha)+\lambda m_i+v(m_i)\big) \\ &\quad= \frac{1}{N}\sum_{i\le N}\Phi_{m_i,\xi}\big(0,\lambda m_i+v(m_i)\big) = \int \Phi_{a,\xi}(0,\lambda a+v(a))\,d\mu_m(a),\end{aligned}$$

which finishes the proof of the first claim. The second claim is obvious because, for $\boldsymbol{\sigma} \in B(m,\varepsilon)$, we have $|R(\widetilde{\boldsymbol{\sigma}},m)| < \varepsilon$. □

Next, we will show that the minimizer in (7.64) is finite under some assumption on μ. The case where this assumption is violated will be handled differently.

LEMMA 7.11. *If* $s = \int a(1-a)d\mu(a) > 0$ *then*

$$\lim_{|\lambda|\to\infty} \int \Phi_{a,\xi}(0, \lambda a + v(a))d\mu(a) = +\infty. \tag{7.71}$$

If $s = \int a(1-a)d\mu(a) = 0$ *then*

$$\inf_{\lambda\in\mathbb{R}} \int \Phi_{a,\xi}(0, \lambda a + v(a))d\mu(a) = \Phi_{0,\xi}(0, v(0))\mu(\{0\}). \tag{7.72}$$

PROOF. Note that the boundary condition in (7.63) satisfies

$$\Phi_{a,\xi}(1-q, x) = \log 2 - ax + \log\cosh(x) \geq -ax + |x|.$$

Moreover, $\Phi_{a,\xi}(0, x)$ is nondecreasing in ξ and, therefore, setting ξ to be identically 0 on $[0, 1-q)$ for the lower bound and letting $g \sim N(0, \xi_q'(1-q))$, using the Feynman-Kac formula we get

$$\begin{aligned}
&\Phi_{a,\xi}(0, \lambda a + v(a)) \\
&\quad \geq \mathbb{E}\big(-a(\lambda a + v(a) + g) + |\lambda a + v(a) + g|\big) \\
&\quad \geq -|\lambda|a^2 - av(a) + |\lambda|a - |v(a)| - \mathbb{E}|g| \geq |\lambda|a(1-a) - c,
\end{aligned} \tag{7.73}$$

for some constant c that depends on v and ξ. Integrating over μ finishes the proof in the case when $s > 0$.

Let us now consider the case when μ is concentrated on $\{0, 1\}$. Then,

$$\int \Phi_{a,\xi}(0, \lambda a + v(a))d\mu(a) = \Phi_{0,\xi}(0, v(0))\mu(\{0\}) + \Phi_{1,\xi}(0, \lambda + v(0))\mu(\{1\}).$$

The first term does not depend on λ and, when $a = 1$, the boundary condition

$$\Phi_{1,\xi}(1-q, x) = \log 2 - x + \log\cosh(x) = \log(1 + e^{-2x})$$

is decreasing. This means that the infimum will be achieved by letting $\lambda \to +\infty$ and, since

$$\lim_{\lambda\to+\infty} \Phi_{1,\xi}(1-q, \lambda + v(1) + x) = 0$$

for all x, by the monotone convergence theorem, $\lim_{\lambda\to+\infty} \Phi_{1,\xi}(0, \lambda + v(1)) = 0$. This proves the second claim. □

PROOF OF PROPOSITION 7.9. First of all, note that $x \mapsto \Phi_{a,\xi}(0, x)$ is a twice differentiable convex function with uniformly bounded first and second derivatives, see Subsection 13.1. As a result, for any $\mu \in \mathcal{M}_{0,1}$,

$$\lambda \mapsto \int \Phi_{a,\xi}(0, \lambda a + v(a))d\mu(a)$$

is twice differentiable with

$$\begin{aligned}
\frac{d}{d\lambda} \int \Phi_{a,\xi}(0, \lambda a + v(a))d\mu(a) &= \int a\partial_x \Phi_{a,\xi}(0, \lambda a + v(a))d\mu(a), \\
\frac{d^2}{d\lambda^2} \int \Phi_{a,\xi}(0, \lambda a + v(a))d\mu(a) &= \int a^2\partial_{xx} \Phi_{a,\xi}(0, \lambda a + v(a))d\mu(a).
\end{aligned}$$

In addition, the second derivative is uniformly bounded over all choices of λ, v, and μ.

We will first consider the case $s = \int a(1-a)d\mu(a) > 0$. By the previous lemma, the infimum in (7.64) is achieved at some finite $\lambda = \lambda_0$, which satisfies

$$\int \partial_x \Phi_{a,\xi}(0, \lambda_0 a + v(a))d\mu(a) = 0. \tag{7.74}$$

By (7.68) and Lemma 7.10, for $\gamma \geq 0$,

$$\Psi_N(m; \lambda_0, B^{\pm}(m,\varepsilon)) \leq \int \Phi_{a,\xi}(0, (\lambda_0 \pm \gamma)a + v(a))d\mu_m(a) - \gamma\varepsilon. \tag{7.75}$$

When $\gamma = 0$, the right hand side equals $\Psi_N(m; \lambda_0)$. Since $\Phi_{a,\xi}(0, (\lambda+\gamma)a+v(a))$ is bounded and continuous in a from (13.4) and $\mu_m \to \mu$, the right hand side converges to

$$\int \Phi_{a,\xi}(0, (\lambda_0 \pm \gamma)a + v(a))d\mu(a) - \gamma\varepsilon.$$

Since the derivatives of convex functions converge to the derivative of the limit, the derivative of the right-hand side of (7.75) in γ at $\gamma = 0$ (denote it D_N) converges to

$$D_N \to \pm \int a\partial_x \Phi_{a,\xi}(0, \lambda_0 a + v(a))d\mu(a) - \varepsilon = -\varepsilon,$$

by (7.74). Finally, from the discussion at beginning of our proof, the second derivative of the right-hand side of (7.75) in γ is bounded uniformly over all parameters by some constant L. Therefore, using Taylor's formula and taking $\gamma = -D_N/L > 0$,

$$\begin{aligned}\Psi_N(m; \lambda_0, B^{\pm}(m,\varepsilon)) &\leq \Psi_N(m; \lambda_0, \Sigma_N) + D_N\gamma + \frac{L\gamma^2}{2} \\ &= \Psi_N(m; \lambda_0, \Sigma_N) - \frac{D_N^2}{2L} \leq \Psi_N(m; \lambda_0, \Sigma_N) - \frac{\varepsilon^2}{4L},\end{aligned} \tag{7.76}$$

for large N.

Let us define a random analogue of (7.67),

$$\psi_N(m; \lambda, A) = \frac{1}{N}\log \sum_{\alpha \in \mathbb{N}^r} v_\alpha \sum_{\sigma \in A} \exp \sum_{i \leq N} \big(g_{\xi'_q,i}(\alpha) + \lambda m_i + v(m_i)\big)\widetilde{\sigma}_i. \tag{7.77}$$

Let us recall (see e.g. Chapter 2 of [44]) that the weights v_α of the Ruelle probability cascades are defined as $w_\alpha / \sum_\alpha w_\alpha$, where w_α are defined as a certain product along the path of the tree of values of Poisson processes. In particular, we can rewrite $\psi_N(m; \lambda, A)$ as

$$\begin{aligned}&R_1 - R_2 := \\ &\frac{1}{N}\log \sum_{\alpha \in \mathbb{N}^r} w_\alpha \sum_{\sigma \in A} \exp \sum_{i \leq N} \big(g_{\xi'_q,i}(\alpha) + \lambda m_i + v(m_i)\big)\widetilde{\sigma}_i - \frac{1}{N}\log \sum_{\alpha \in \mathbb{N}^r} w_\alpha.\end{aligned}$$

By the Bolthausen-Sznitman invariance property [44, theorem 2.6] for the Poisson processes Π_t on $(0,\infty)$ with the mean measure $tx^{-1-t}dx$ for $t \in (0,1)$, both terms above are equal in distribution (see the proof of [44, theorem 2.9]) to

$$R_j = c_j + \frac{1}{N}\log \sum_{x\in\Pi_{\zeta_0}} x,$$

for some constants c_1, c_2 (note, however, that the two terms are not independent of each other). This implies that

$$\mathbb{P}\left(\left|\psi_N(m;\lambda,A) - \Psi_N(m;\lambda,A)\right| \geq \frac{2t}{N}\right) \leq 2\mathbb{P}\left(\left|\log \sum_{x\in\Pi_{\zeta_0}} x - \mathbb{E}\log \sum_{x\in\Pi_{\zeta_0}} x\right| \geq t\right).$$

In other words, the fluctuations of $\psi_N(m;\lambda,A)$ are of order $1/N$. The bound in (7.76) implies that, with high probability,

$$\psi_N(m;\lambda_0, B^{\pm}(m,\varepsilon)) \leq \psi_N(m;\lambda_0,\Sigma_N) - \frac{\varepsilon^2}{8L}.$$

Since Σ_N is a union of $B(m,\varepsilon)$, $B^+(m,\varepsilon)$ and $B^-(m,\varepsilon)$, this forces that

$$\psi_N(m;\lambda_0, B^+(m,\varepsilon)\cup B^-(m,\varepsilon)) \leq \frac{\log 2}{N} + \psi_N(m;\lambda_0,\Sigma_N) - \frac{\varepsilon^2}{8L},$$

where the right-hand side used the bound $\log(a+b) \leq \log 2 + \max(\log a, \log b)$ for $a, b > 0$. Consequently,

$$\psi_N(m;\lambda_0,\Sigma_N) + \frac{1}{N}\log\left(1 - 2e^{-\varepsilon^2 N/(8L)}\right) \leq \psi_N(m;\lambda_0, B(m,\varepsilon)) \leq \psi_N(m;\lambda_0,\Sigma_N),$$

where the left-hand side used the fact that if $\frac{1}{N}\log(a+b) \leq \frac{1}{N}\log(a+b+c) - \delta$ for some $a, b, c, \delta > 0$, then $(a+b)/(a+b+c) \leq e^{-N\delta}$ so that

$$\frac{1}{N}\log(a+b+c) + \frac{1}{N}\log(1-e^{-\delta N}) \leq \frac{1}{N}\log c.$$

Since λ_0 was the minimizer, by Lemma 7.10 we get

$$\lim_{N\to\infty} \psi_N(m;\lambda_0, B(m,\varepsilon)) = \inf_{\lambda\in\mathbb{R}} \int \Phi_{a,\xi}(0, \lambda a + v(a))d\mu(a). \tag{7.78}$$

Finally, using the second claim in Lemma 7.10 and letting $\varepsilon \to 0$ finishes the proof.

It remains to consider the case when μ is concentrated on $\{0,1\}$. By Lemma 7.6, to compute the limit of $\Psi_N(m,\varepsilon)$, we can choose any sequence $m = m^N$ such that $\mu_m \to \mu$. In particular, we can choose μ_m also concentrated on $\{0,1\}$. Let us suppose that the first N_1 coordinates are $m_i = 0$ and the last $N - N_1$ coordinates are $m_i = 1$, and $N_1/N \to \mu(\{0\})$.

In this case, the condition $\boldsymbol{\sigma} \in B(m, \varepsilon)$, or $|R(\widetilde{\boldsymbol{\sigma}}, m)| < \varepsilon$, means that

$$2\sum_{i=1}^{N} \mathbf{1}(m_i = 1, \sigma_i = -1) = (m \cdot m - m \cdot \sigma) < N\varepsilon.$$

This means that, when $m_i = 0$, there are no constraints on σ_i and, when $m_i = 1$, we must have $\sigma_i = 1$ with at most $N\varepsilon/2$ exceptions. This means that in the definition of $\Psi_N(m, \varepsilon)$, we can replace the sum in the exponent by $\sum_{i \le N_1} (g_{\xi'_q, i}(\alpha) + v(0))\sigma_i$ (for example, using interpolation) and we can replace the constraint $\boldsymbol{\sigma} \in B(m, \varepsilon)$ by $(\sigma_1, \ldots, \sigma_{N_1}) \in \Sigma_{N_1}$. This will change $\Psi_N(m, \varepsilon)$ by at most $O(\varepsilon \log \varepsilon^{-1})$. This implies that

$$\begin{aligned}\Psi_N(m, \varepsilon) &\approx \frac{1}{N}\mathbb{E} \log \sum_{\alpha \in \mathbb{N}^r} v_\alpha \sum_{\Sigma_{N_1}} \exp \sum_{i \le N_1} (g_{\xi'_q, i}(\alpha) + v(0))\widetilde{\sigma}_i \\ &= \frac{N_1}{N}\mathbb{E} \log \sum_{\alpha \in \mathbb{N}^r} v_\alpha 2\cosh\big(g_{\xi'_q}(\alpha) + v(0)\big) = \frac{N_1}{N}\Phi_{0,\xi}(0, v(0)).\end{aligned}$$

Comparing this with the second claim in Lemma 7.11 finishes the proof in the case when $s = 0$. □

8 Properties of TAP representations

We present the proof of Theorem 3.4. To prepare for our proof, Subsection 8.1 first establishes connections among the Parisi PDEs introduced in the introduction, while Subsection 8.2 derives some uniform upper bounds for $\Phi_{a,\zeta}(0, v(a))$. The details of the proof of Theorem 3.4 are given in Subsection 8.3.

8.1 Relations among Parisi PDEs

Recall that the Parisi PDE Φ_ζ defined in (2.6) has the boundary condition

$$\Phi_\zeta(1, x) = \log 2 + \log\cosh x.$$

For $a \in \mathbb{R}$ and $\zeta \in \mathcal{M}_{0,1}$, denote by $\Theta_{a,\zeta}(t, x)$ the solution of the Parisi PDE on $[0, 1] \times \mathbb{R}$,

$$\partial_t \Theta_{a,\zeta} = -\frac{\xi''(t)}{2}\big(\partial_{xx}\Theta_{a,\zeta} + \zeta(t)\big(\partial_x \Theta_{a,\zeta}\big)^2\big) \tag{8.1}$$

with the boundary condition

$$\Theta_{a,\zeta}(1, x) = \log 2 - ax + \log\cosh x = \log \sum_{\sigma = \pm 1} e^{(\sigma - a)x}. \tag{8.2}$$

In other words, the two solutions Φ_ζ and $\Theta_{a,\zeta}$ satisfy the same PDE, but with different boundary conditions. The following proposition shows that these two solutions are indeed connected through an elementary formula.

PROPOSITION 8.1. *For any* $x, a \in \mathbb{R}$, $t \in [0,1]$, *and* $\zeta \in \mathcal{M}_{0,1}$,

$$\Theta_{a,\xi}(t,x) = \Phi_\zeta\left(t, x - a\int_t^1 \xi''(s)\zeta(s)ds\right) - ax + \frac{a^2}{2}\int_t^1 \xi''(s)\zeta(s)ds. \tag{8.3}$$

Remark 8.2. Recall $\Phi_{a,\xi}$ from (3.24). As an immediate consequence of this proposition, by shifting the measure ζ from $[0, 1-q]$ to $[q, 1]$, we can now unify all Parisi PDEs together. To see this, let $a \in [-1,1]$ and $q \in [0,1]$. For any $\zeta \in \mathcal{M}_{0,1}$, if we recall the shift operator θ_q in (3.37) and let $\zeta_q = \theta_q\zeta \in \mathcal{M}_{0,1-q}$, defined by

$$\zeta_q(t) = \zeta(q+t), \quad \forall t \in [0, 1-q],$$

then, for any $x \in \mathbb{R}$,

(8.4)

$$\begin{aligned}\Phi_{a,\zeta_q}(0,x) &= \Theta_{a,\xi}(q,x)\\ &= \Phi_\zeta\left(q, x - a\int_q^1 \xi''(s)\zeta(s)ds\right) - ax + \frac{a^2}{2}\int_q^1 \xi''(s)\zeta(s)ds.\end{aligned}$$

Note that since Φ_{a,ζ_q} attains the global minimum at $\Psi(a,\zeta_q)$, taking infimum over x leads to

$$\Phi_{a,\zeta_q}(0,\Psi(a,\zeta_q)) = \Lambda_\xi(q,a) - \frac{a^2}{2}\int_q^1 \xi''(s)\zeta(s)ds, \quad \forall a \in (-1,1). \tag{8.5}$$

PROOF OF PROPOSITION 8.1. It suffices to prove (8.3) only for continuous ζ. Denote the right hand side of (8.3) by $f(t,x)$ and set

$$b(t,x) = x - a\int_t^1 \xi''(s)\zeta(s)ds.$$

Then

$$\partial_t f(t,x) = \partial_t\Phi_\zeta(t,b(t,x)) + a\xi''(t)\zeta(t)\partial_x\Phi_\zeta(t,b(t,x)) - \frac{a^2}{2}\xi''(t)\zeta(t)$$

and

$$\begin{aligned}\partial_x f(t,x) &= \partial_x\Phi_\zeta(t,b(t,x)) - a,\\ \partial_{xx} f(t,x) &= \partial_{xx}\Phi_\zeta(t,b(t,x)).\end{aligned}$$

From this, a direct verification gives

$$\partial_t f(t,x) = -\frac{\xi''(t)}{2}\big(\partial_{xx} f(t,x) + \zeta(t)\big(\partial_x f(t,x)\big)^2\big).$$

Note that $f(1,x) = \log 2 - ax + \log\cosh x = \Theta_{a,\xi}(1,x)$. Finally, we recall that it was proved in Jagannath-Tobasco [32, Lemma 13] that the classical Parisi PDE has a unique solution. The same proof therein applies to the current setting with no essential changes and yields the uniqueness of the Parisi PDE solution with boundary condition $\log 2 - ax + \log\cosh x$. □

8.2 Uniform upper bounds for $\Phi_{a,\xi}(0, v(a))$

Note that

$$\sup_{a\in[0,1)} (1-a)\tanh^{-1}(a) \le 1, \text{ and } \lim_{a\uparrow 1}(1-a)\tanh^{-1}(a) = 0.$$

As a result, for any $v \in V \cup \overline{V}$, $(1-a)v(a)$ is bounded from above by some absolute constant. Since $\Phi_{a,\xi}$ is nondecreasing in $\zeta \in \mathcal{M}_{0,1-q}$, comparing with $\zeta \equiv 1$, one can see that, with $c^2 = \xi_q'(1-q) \le \xi'(1)$,

$$\begin{aligned}\Phi_{a,\xi}(0,x) &\le \log \sum_{\sigma=\pm 1} \exp\big((\sigma-1)x + \frac{1}{2}(\sigma-a)^2c^2\big)\\ &= (1-a)x + \frac{1}{2}(1-a)^2c^2 + \log\big(1+e^{2a(c^2-x)}\big).\end{aligned}$$

In particular, for some constant c_ξ, for all $v \in V \cup \overline{V}$,

(8.6)
$$\Phi_{a,\xi}(0, v(a)) \le (1-a)v(a) + \frac{1}{2}(1-a)^2c^2 + \log\big(1+e^{2a(c^2-v(a))}\big) \le c_\xi.$$

Moreover, if we take $v(a) = v_{\zeta'}(a) = \Psi(a,\zeta')$ for some $\zeta' \in \mathcal{M}_{q'}$ and use Lemma 13.3 below, which states that

$$c_1' + c_2'\tanh^{-1}(a) \le v_{\zeta'}(a) \le c_1 + c_2\tanh^{-1}(a)$$

for some absolute constants $c_1, c_1' \in \mathbb{R}$, $c_2, c_2' > 0$, we also have

$$\begin{aligned}0 \le \Phi_{a,\xi}(0, v_{\zeta'}(a)) &\le (1-a)(c_1 + c_2\tanh^{-1}(a)) + \frac{1}{2}(1-a)^2c^2\\ &\quad + \log\big(1+\exp\big(2ac^2 - 2c_1'a - 2c_2'a\tanh^{-1}(a)\big)\big) =: M(a),\end{aligned} \tag{8.7}$$

We can see that $\lim_{a\uparrow 1} M(a) = 0$, which shows that $\Phi_{a,\xi}(0, v_{\zeta'}(a))$ is small in the neighbourhood of 1 uniformly over the choice of ζ and ζ'. Using (8.5), this implies that

$$0 \le \Lambda_\zeta(q,a) - \frac{a^2}{2}\int_q^1 \xi''(s)\zeta(s)ds \le M(|a|), \tag{8.8}$$

so the expression in the middle goes to zero as $a \to \pm 1$, uniformly over q and ζ.

8.3 Proof of Theorem 3.4

Our approach relies on the convexity of the Parisi functional $\overline{\mathcal{P}}_\mu^v(\cdot,\cdot)$ on the band (defined in (3.35)) as well as some computations on the directional derivative of this functional. Let μ be a probability measure on $[0,1]$.

PROOF OF (1). First we establish the first equality in (i). To see this, observe that (8.5) and (8.7) together imply that whenever $\zeta \in \mathcal{M}_{0,1}$ and $\zeta_q = \theta_q\zeta \in$

$\mathcal{M}_{0,1-q}$ satisfy $\zeta_q(t) = \zeta(q+t)$ for all $t \in [0, 1-q]$, we have

$$\Lambda_\zeta(q,1) = \lim_{a\to 1} \Lambda_\zeta(q,a) = \frac{1}{2}\int_q^1 \xi''(s)\zeta(s)ds.$$

Using (8.5) and this equation yield that

$$\begin{aligned}
&\overline{\mathcal{P}}_\mu^{v_{\zeta_q}}(0,\zeta_q)\\
&= \int_{[0,1)} \left(\Lambda_\zeta(q,a) - \frac{a^2}{2}\int_q^1 \xi''(s)\zeta(s)ds\right)d\mu(a) - \frac{1}{2}\int_q^1 (s-q)\xi''(s)\zeta(s)ds\\
&= \int_{[0,1]} \left(\Lambda_\zeta(q,a) - \frac{a^2}{2}\int_q^1 \xi''(s)\zeta(s)ds\right)d\mu(a) - \frac{1}{2}\int_q^1 (s-q)\xi''(s)\zeta(s)ds\\
&= \mathrm{TAP}(\mu,\zeta),
\end{aligned}$$

which implies that

$$\begin{aligned}
\inf_{\zeta\in\mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_\zeta}(0,\zeta) &= \inf_{\zeta\in\mathcal{M}_{0,1}} \overline{\mathcal{P}}_\mu^{v_{\zeta_q}}(0,\zeta_q)\\
&= \inf_{\zeta\in\mathcal{M}_{0,1}} \mathrm{TAP}(\mu,\zeta) = \inf_{\zeta\in\mathcal{M}_{q,1}} \mathrm{TAP}(\mu,\zeta) = \mathrm{TAP}(\mu).
\end{aligned}$$

This establishes one of the equalities in (i). For the rest of the equalities, they follow immediately if the following claim is valid: for any $\zeta \in \mathcal{M}_{0,1-q}$,

$$\begin{aligned}
\inf_{v\in V}\int_{[0,1]} \Phi_{a,\zeta}(0,v(a))d\mu(a) &= \int_{[0,1)} \Phi_{a,\zeta}(0,\Psi(a,\zeta))d\mu(a)\\
&= \inf_{v\in\overline{V}}\int_{[0,1)} \Phi_{a,\zeta}(0,v(a))d\mu(a).
\end{aligned} \tag{8.9}$$

To show this claim, observe that for any $a \in [0,1)$, $\Phi_{a,\zeta}(0,\cdot)$ is a strictly convex function (see (13.6)) and has a unique global minimum at $\Psi(a,\zeta)$ since

$$\partial_x \Phi_{a,\zeta}(a,\Psi(a,\zeta)) = 0.$$

These imply that

$$\Phi_{a,\zeta}(a,\Psi(a,\zeta)) \le \Phi_{a,\zeta}(a,x), \quad \forall x \in \mathbb{R}. \tag{8.10}$$

From this, Lemma 13.3, and noting that $\Phi_{a,\zeta}$ is always nonnegative, we have

$$\begin{aligned}
\int_{[0,1)} \Phi_{a,\zeta}(0,\Psi(a,\zeta))d\mu(a) &\le \inf_{v\in V}\int_{[0,1]} \Phi_{a,\zeta}(0,v(a))d\mu(a),\\
\int_{[0,1)} \Phi_{a,\zeta}(0,\Psi(a,\zeta))d\mu(a) &= \inf_{v\in\overline{V}}\int_{[0,1)} \Phi_{a,\zeta}(0,v(a))d\mu(a).
\end{aligned} \tag{8.11}$$

The second line here gives the second equality of (8.9). To show the first equality of (8.9), it remains to establish the reverse inequality for the first equation of (8.11).

To this end, for $L > 0$, set $v_L \in V$ by $v_L(a) = \Psi(a,\zeta) \wedge L$ for $a \in [0,1)$ and $v_L(1) = L$. Write

$$\int_{[0,1]} \Phi_{a,\xi}(0, v_L(a))d\mu(a) = \int_{[0,1)} \Phi_{a,\xi}(0, v_L(a))d\mu(a) + \Phi_{1,\xi}(0,L)\mu(\{1\}).$$

Passing to the limit via the bound (8.6) and the dominated convergence theorem gives that

$$\begin{aligned}&\limsup_{L\to\infty} \int_{[0,1]} \Phi_{a,\xi}(0, v_L(a))d\mu(a)\\ &\le \int_{[0,1)} \Phi_{a,\xi}(0, \Psi(a,\zeta))d\mu(a) + \limsup_{L\to\infty} \Phi_{1,\xi}(0,L)\mu(\{1\}).\end{aligned}$$

Note that $\Phi_{1,\xi}(1,x) = \log(1+e^{-2x})$ is a strictly decreasing function with $\Phi_{1,\xi}(1,\infty) = 0$. These properties are also valid for $\Phi_{1,\xi}(0,\cdot)$, which can be seen from the representation (13.2). Hence, $\limsup_{L\to\infty} \Phi_{1,\xi}(0,L) = 0$ and consequently,

$$\begin{aligned}\inf_{v\in V} \int_{[0,1]} \Phi_{a,\xi}(0, v(a))d\mu(a) &\le \limsup_{L\to\infty} \int_{[0,1]} \Phi_{a,\xi}(0, v_L(a))d\mu(a)\\ &\le \int_{[0,1)} \Phi_{a,\xi}(0, \Psi(a,\zeta))d\mu(a).\end{aligned}$$

This completes the proof of our claim. □

PROOF OF (II). By part (i), we can work with any of the four representations. We will use the first one, TAP(μ). If we denote

$$D_\xi(q,a) := \Lambda_\xi(q,a) - \frac{a^2}{2}\int_q^1 \xi''(s)\zeta(s)ds$$

then, for μ with $\int a^2\, d\mu(a) = q$, we can rewrite

$$\mathrm{TAP}(\mu,\zeta) := \int D_\xi(q,a)d\mu(a) - \frac{1}{2}\int_q^1 (s-q)\xi''(s)\zeta(s)ds.$$

By (8.8), for any $\varepsilon > 0$, we can find $\eta \in (0,1)$ such that, $0 \le D_\xi(q,a) \le \varepsilon$ for $|a| \ge \eta$. Since, $\Lambda_\xi(q,a)$ and $D_\xi(q,a)$ are even (and concave), if we let

$$\mathrm{TAP}^\eta(\mu,\zeta) := \int D_\xi(q, |a|\wedge\eta)d\mu(a) - \frac{1}{2}\int_q^1 (s-q)\xi''(s)\zeta(s)ds$$

then

$$\left|\mathrm{TAP}^\eta(\mu,\zeta) - \mathrm{TAP}(\mu,\zeta)\right| \le \varepsilon. \tag{8.12}$$

Using the fact that $\partial_x \Phi_\xi(q,\cdot)$ is strictly increasing with $\partial_x\Phi_\xi(q,\pm\infty) = \pm 1$, it is easy to check that, for $|a| < \eta$, the infimum in $\Lambda_\xi(q,a) = \inf_{x\in\mathbb{R}}(\Phi_\xi(q,x) - ax)$ is achieved on $x \in [-L, L]$, where L depends on η only. This implies that $\Lambda_\xi(q,a)$

and $D_\zeta(q,a)$ are Lipschitz on $[-\eta,\eta]$ and $D_\zeta(q,|a|\wedge\eta)$ is Lipschitz on $[-1,1]$, uniformly over q and ζ. Therefore,

$$\sup_{q,\zeta}\left|\int D_\zeta(q,|a|\wedge\eta)d\mu(a)-\int D_\zeta(q,|a|\wedge\eta)d\mu'(a)\right|\le Cd_{\mathrm{BL}}(\mu,\mu'),$$

where $d_{\mathrm{BL}}(\mu,\mu')$ is the bounded Lipschitz metric on $\mathcal{M}_{0,1}$. Moreover, since $\Phi_\zeta(q,x)$ is Lipschitz in q and d_1-Lipschitz in ζ (with Lipschitz constants that depend only on ξ), these properties are inherited by $\Lambda_\zeta(q,a)$, $D_\zeta(q,a)$ and $D_\zeta(q,|a|\wedge\eta)$, and, therefore, $\mathrm{TAP}^\eta(\mu,\zeta)$. If $\mu\to\mu_0$ then $q=\int a^2\,d\mu(a)\to q_0=\int a^2\,d\mu_0(a)$, and all the properties above imply that $\inf_\zeta \mathrm{TAP}^\eta(\mu,\zeta)$ is continuous in μ. By (8.12), this proves that $\mathrm{TAP}(\mu)$ is continuous in μ. □

From now on we can assume that $\mu\neq\delta_1$, because all the remaining claims are trivial in this case.

PROOF OF (III). Let $\zeta_n\in\mathcal{M}_{0,1-q}$ be a sequence that weakly converges to some ζ_0 and satisfies

$$\inf_{\zeta\in\mathcal{M}_{0,1-q}}\overline{\mathcal{P}}_\mu^{v_\zeta}(0,\zeta)=\lim_{n\to\infty}\overline{\mathcal{P}}_\mu^{v_{\zeta_n}}(0,\zeta_n).$$

For any $\delta\in(0,1)$, write

$$\begin{aligned}\int_{[0,1)}\Phi_{a,\zeta_n}(0,\Psi(a,\zeta_n))d\mu(a)&=\int_{[0,\delta]}\Phi_{a,\zeta_n}(0,\Psi(a,\zeta_n))d\mu(a)\\&\quad+\int_{(\delta,1)}\Phi_{a,\zeta_n}(0,\Psi(a,\zeta_n))d\mu(a).\end{aligned}$$

Here, the first term converges to

$$\int_{[0,\delta]}\Phi_{a,\zeta_0}(0,\Psi(a,\zeta_0))d\mu(a)$$

as a consequence of Lemma 13.1 and the fact that $\Phi_{a,\zeta_n}(t,x)$ converges to $\Phi_{a,\zeta_0}(t,x)$ uniformly over all $a\in[0,1], t\in[0,1-q], x\in\mathbb{R}$. As for the second term, note that the inequality (8.7) implies

$$0\le\int_{(\delta,1)}\Phi_{a,\zeta_n}(0,\Psi(a,\zeta_n))d\mu(a)\le\sup_{\delta<a<1}M(a)\to 0$$

as $\delta\uparrow 1$. Hence, we arrive at

$$\lim_{n\to\infty}\int_{[0,1)}\Phi_{a,\zeta_n}(0,\Psi(a,\zeta_n))d\mu(a)=\int_{[0,1)}\Phi_{a,\zeta_0}(0,\Psi(a,\zeta_0))d\mu(a)$$

and

$$\lim_{n\to\infty}\overline{\mathcal{P}}_\mu^{v_{\zeta_n}}(0,\zeta_n)=\overline{\mathcal{P}}_\mu^{v_{\zeta_0}}(0,\zeta_0),$$

which shows that ζ_0 is a minimizer of $\inf_{\zeta\in\mathcal{M}_{0,1-q}}\overline{\mathcal{P}}_\mu^{v_\zeta}(0,\zeta)$.

Next, we establish the uniqueness of ζ_0. Assume that ζ_1 is another minimizer and $\zeta_1 \not\equiv \zeta_0$. For $b \in (0,1)$, let

$$\zeta_b = (1-b)\zeta_0 + b\zeta_1, \quad v_b = (1-b)v_{\zeta_0} + bv_{\zeta_1}.$$

Since $\mu \neq \delta_1$, we can use the strict convexity in (13.7) below to get

$$\int_{[0,1)} \Phi_{a,\zeta_b}(0, v_b(a))d\mu(a) < (1-b)\int_{[0,1)} \Phi_{a,\zeta_0}(0, v_{\zeta_0}(a))d\mu(a) + b\int_{[0,1)} \Phi_{a,\zeta_1}(0, v_{\zeta_1}(a))d\mu(a).$$

On the other hand, note that, by (8.10),

$$\int_{[0,1)} \Phi_{a,\zeta_b}(0, \Psi(a,\zeta_b))d\mu(a) \leq \int_{[0,1)} \Phi_{a,\zeta_b}(0, v_b(a))d\mu(a).$$

This and the above inequality together lead to a contradiction,

(8.13)
$$\inf_{\zeta\in\mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_\zeta}(0,\zeta) < (1-b)\inf_{\zeta\in\mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_\zeta}(0,\zeta) + b\inf_{\zeta\in\mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_\zeta}(0,\zeta) = \inf_{\zeta\in\mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_\zeta}(0,\zeta).$$

Hence, the minimizer must be unique when $\mu \neq \delta_1$. □

PROOF OF (iv). Note that the minimality of ζ_0 implies

$$\inf_{\zeta\in\mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_\zeta}(0,\zeta) = \overline{\mathcal{P}}_\mu^{v_{\zeta_0}}(0,\zeta_0) \geq \inf_{\zeta\in\mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_{\zeta_0}}(0,\zeta) \geq \inf_{\lambda\in\mathbb{R},\zeta\in\mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_{\zeta_0}}(\lambda,\zeta) = \overline{\mathcal{P}}_\mu^{v_{\zeta_0}}.$$

Also note that from (8.10), for any $\zeta \in \mathcal{M}_{0,1-q}$ and $\lambda \in \mathbb{R}$,

$$\overline{\mathcal{P}}_\mu^{v_\zeta}(0,\zeta) \leq \overline{\mathcal{P}}_\mu^{v_{\zeta_0}}(\lambda,\zeta),$$

which leads to

$$\overline{\mathcal{P}}_\mu^{v_{\zeta_0}}(0,\zeta_0) = \inf_{\zeta\in\mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_\zeta}(0,\zeta) \leq \inf_{\lambda\in\mathbb{R},\zeta\in\mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_{\zeta_0}}(\lambda,\zeta) = \overline{\mathcal{P}}_\mu^{v_{\zeta_0}}.$$

This completes our proof. □

PROOF OF (v). Let $\delta_1 \neq \mu \in \mathcal{M}_{0,1}$. Let ζ_0 be the minimizer from part (iii). By part (iv), the pair $(0,\zeta_0)$ is a minimizer of $\overline{\mathcal{P}}_\mu^{v_{\zeta_0}}(\lambda,\zeta)$. Also, by the definition of v_{ζ_0} in (3.31) and (3.26), we have $\partial_x \Phi_{a,\zeta_0}(0, v_{\zeta_0}(a)) = 0$ for all $a \in [0,1)$.

We follow a similar argument as [15, theorem 1] (see also [48, Lemma 4.14]). Let c be the smallest point in the support of ζ_0. Assume on the contrary that $c > 0$. Note that from the optimality of ζ_0 in $\overline{\mathcal{P}}_\mu^{v_{\zeta_0}}(0,\cdot)$, Remark 13.7 below states that

(8.14)
$$\int_{[0,1)} \mathbb{E}\big(\partial_x \Phi_{a,\zeta_0}(c, v_{\zeta_0}(a) + z(c))\big)^2 \, d\mu(a) = c,$$

$$\xi_q''(c)\int_{[0,1)}\mathbb{E}\big(\partial_{xx}\Phi_{a,\xi_0}(c,v_{\xi_0}(a)+z(c))\big)^2\,d\mu(a)\le 1, \tag{8.15}$$

where $z(c)$ is a centered normal random variable with variance $\xi_q'(c)$. Define an auxiliary function $A:[0,c]\to[0,c]$ by

$$A(t)=\int_{[0,1)}\mathbb{E}\big(\partial_x\Phi_{a,\xi_0}(c,v_{\xi_0}(a)+z_1(t)\big)\big(\partial_x\Phi_{a,\xi_0}(c,v_{\xi_0}(a)+z_2(t)\big)d\mu(a),$$

where $z_1(t)$ and $z_2(t)$ are jointly Gaussian random variables with mean zero and variance $\mathbb{E}(z_1(t))^2=\mathbb{E}(z_2(t))^2=\xi_q'(c)$ and $\mathbb{E}z_1(t)z_2(t)=\xi_q'(t)$. From this construction and (8.14), evidently $A(c)=c$. Since $\partial_x\Phi_{a,\xi_0}(0,v_{\xi_0}(a))\equiv 0$, we also have that

$$\begin{aligned}A(0)&=\int_{[0,1)}\big(\mathbb{E}\partial_x\Phi_{a,\xi_0}(c,v_{\xi_0}(a)+z(c))\big)^2\,d\mu(a)\\&=\int_{[0,1)}\big(\mathbb{E}\partial_x\Phi_{a,\xi_0}(0,v_{\xi_0}(a))\big)^2\,d\mu(a)=0,\end{aligned}$$

where the second equality holds because $\zeta_0(s)=0$ for $s\in[0,c)$ and

$$\Phi_{a,\xi_0}(0,x)=\mathbb{E}\Phi_{a,\xi_0}(c,x+z(c)).$$

Next, a direct differentiation using Gaussian integration by parts, the bounds in (13.1), and the dominated convergence theorem gives

$$A'(t)=$$
$$\xi_q''(t)\int_{[0,1)}\mathbb{E}\big(\partial_{xx}\Phi_{a,\xi_0}(c,v_{\xi_0}(a)+z_1(t)\big)\big(\partial_{xx}\Phi_{a,\xi_0}(c,v_{\xi_0}(a)+z_2(t)\big)d\mu(a),$$

from which we see that

$$0\le A'(t)<A'(c)\le 1$$

for all $t\in[0,c)$, where the first inequality is obtained by using conditional expectation and integrating the independent components of $z_1(t)$ and $z_2(t)$ first, second inequality is by the Cauchy-Schwarz inequality, and third inequality follows by (8.15). This contradicts that both $A(0)=0$ and $A(c)=c$. Hence, the smallest point in the support of ζ_0 must be zero and this completes our proof. □

PROOF OF (vi). We show that if $\overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(0,\zeta_1)=\inf_{\lambda,\zeta}\overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(\lambda,\zeta)$ then $\zeta_1=\zeta_0$. It suffices to show that

$$\overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(0,\zeta_1)=\overline{\mathbf{T}}(\mu)=\inf_{v\in\overline{V},\,\zeta\in\mathcal{M}_{0,1-q}}\overline{\mathcal{P}}_\mu^{v}(0,\zeta). \tag{8.16}$$

Indeed, if this holds then, from part (i),

$$\overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(0,\zeta_1)=\inf_{\zeta\in\mathcal{M}_{0,1-q}}\overline{\mathcal{P}}_\mu^{v_\zeta}(0,\zeta).$$

Then, by part (iii) and the assumption that $\mu \neq \delta_1$, we get $\zeta_1 = \zeta_0$. To prove (8.16), we can argue as follows. Notice that from (13.7),

$$(v, \zeta) \in \overline{V} \times \mathcal{M}_{0,1-q} \mapsto \overline{\mathcal{P}}_\mu^v(0, \zeta)$$

is a convex function. From the minimality of $(0, \zeta_1)$ in $\overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(\lambda, \zeta)$,

$$\overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(0, \zeta_1) \leq \overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(0, \zeta), \quad \forall \zeta \in \mathcal{M}_{0,1-q}$$

and, by (8.10),

$$\overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(0, \zeta_1) \leq \overline{\mathcal{P}}_\mu^{v}(0, \zeta_1), \quad \forall v \in \overline{V}.$$

In other words, (v_{ζ_1}, ζ_1) is a local minimum of $\overline{\mathcal{P}}_\mu^v(0, \zeta)$ in the two coordinates v and ζ separately. From Lemma 13.6 below, for any $\zeta \in \mathcal{M}_{0,1-q}$ and $v \in \overline{V}$ satisfying $v = v_{\zeta_1}$ on $[\delta, 1)$ for some $\delta \in (0, 1)$, the directional derivative of $(v, \zeta) \to \overline{\mathcal{P}}_\mu^v(0, \zeta)$ exists along the direction from (v_{ζ_1}, ζ_1) to (v, ζ) and is equal to (13.18). From this, it can be checked that

$$\begin{aligned}
&\frac{d}{db}\overline{\mathcal{P}}_\mu^{(1-b)v_{\zeta_1}+bv}(0, (1-b)\zeta_1 + b\zeta)\Big|_{b=0^+} \\
&= \frac{d}{db}\overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(0, (1-b)\zeta_1 + b\zeta)\Big|_{b=0^+} + \frac{d}{db}\overline{\mathcal{P}}_\mu^{(1-b)v_{\zeta_1}+bv}(0, \zeta_1)\Big|_{b=0^+} \geq 0,
\end{aligned}$$

where the last inequality is a consequence of the previous two displays. With this, for any $\varepsilon > 0$, there exists some small $b > 0$ such that

$$\overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(0, \zeta_1) - b\varepsilon \leq \overline{\mathcal{P}}_\mu^{(1-b)v_{\zeta_1}+bv}(0, (1-b)\zeta_1 + b\zeta).$$

On the other hand, using the convexity on the right-hand side yields

$$\overline{\mathcal{P}}_\mu^{(1-b)v_{\zeta_1}+bv}(0, (1-b)\zeta_1 + b\zeta) \leq (1-b)\overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(0, \zeta_1) + b\overline{\mathcal{P}}_\mu^{v}(0, \zeta).$$

Putting these two inequalities together gives $\overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(0, \zeta_1) \leq \varepsilon + \overline{\mathcal{P}}_\mu^{v}(0, \zeta)$, and letting $\varepsilon \downarrow 0$, we get

$$\overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(0, \zeta_1) \leq \mathcal{P}_\mu^{v}(0, \zeta). \tag{8.17}$$

Note that we proved this for $v \in \overline{V}$ satisfying $v = v_{\zeta_1}$ on $[\delta, 1)$ for any $\delta \in (0, 1)$. In what follows, we show that this implies the same inequality for all $v \in V$ and $\zeta \in \mathcal{M}_{0,1-q}$.

For any $\delta \in (1/3, 1)$, let $a(\delta) := (3\delta - 1)/2$. Then $a(\delta) \in (0, \delta)$. For any $v \in V$, since v is bounded on $[0, 1]$, we can construct $v_\delta \in \overline{V}$ so that $v_\delta = v$ on $[0, (3\delta - 1)/2]$ and $v_\delta = v_{\zeta_1}$ on $[\delta, 1)$ as long as δ is sufficiently close to 1. From

this, write

$$\begin{aligned}&\int_{[0,1)} \Phi_{a,\xi}(0, v_\delta(a))\mu(da) - \int_{[0,a(\delta))} \Phi_{a,\xi}(0, v(a))\mu(da)\\ &= \int_{[a(\delta),1)} \Phi_{a,\xi}(0, v_\delta(a))\mu(da)\\ &= \int_{[a(\delta),\delta)} \Phi_{a,\xi}(0, v_\delta(a))\mu(da) + \int_{[\delta,1)} \Phi_{a,\xi}(0, v_{\zeta_1}(a))\mu(da)\end{aligned}$$

and use the bounds (8.6) and (8.7) to get

$$\begin{aligned}&\left|\int_{[0,1)} \Phi_{a,\xi}(0, v_\delta(a))\mu(da) - \int_{[0,a(\delta))} \Phi_{a,\xi}(0, v(a))\mu(da)\right|\\ &\le c_\xi \mu([a(\delta), \delta)) + M(\delta),\end{aligned}$$

Here the second term vanishes as $\delta \to 1^-$. The first term can be handled as follows. Note that $\mu([s, 1))$ is a nonincreasing function, so $\lim_{s\to 1^-} \mu([s, 1))$ exists. This implies that

$$\lim_{\delta\to 1^-} \mu([a(\delta), \delta)) = \lim_{\delta\to 1^-} \mu([a(\delta), 1)) - \lim_{\delta\to 1^-} \mu([\delta, 1)) = 0.$$

Hence,

$$\lim_{\delta\to 1^-} \left|\int_{[0,1)} \Phi_{a,\xi}(0, v_\delta(a))\mu(da) - \int_{[0,a(\delta))} \Phi_{a,\xi}(0, v(a))\mu(da)\right| = 0$$

and, starting with (8.17) for v_δ,

$$\overline{\mathcal{P}}_\mu^{v_{\zeta_1}}(0, \zeta_1) \le \lim_{\delta\to 1^-} \overline{\mathcal{P}}_\mu^{v_\delta}(0, \zeta) = \overline{\mathcal{P}}_\mu^{v'}(0, \zeta) \le \mathcal{P}_\mu^{v}(0, \zeta),$$

where v' is the restriction of v on $[0, 1)$. This establishes (8.17) for all $v \in V$. Now from (i), we see that ζ_1 minimizes

$$\inf_{\zeta\in\mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^{v_\zeta}(0, \zeta)$$

and from (iii), $\zeta_0 = \zeta_1$. This finishes the proof. □

9 Optimizing over the External Field

In this section, we will prove the upper and lower bounds of Lemmas 4.1 and 4.2 on the limiting replicated free energy on the band. We will use the representation $\mathbf{T}(\mu)$ of (3.36), which by Theorem 3.4 (i) is equivalent to TAP(μ). The upper bound of Lemma 4.1 will be straightforward to prove, by introducing an arbitrary continuous external field and then applying the Guerra upper bound. The lower bound contains the key step, where will need to use the optimal external field found in the last section (for which the Parisi measure has zero in the support).

9.1 Proof of Lemma 4.1

Notice that, since

$$n\mathbb{E}\,\mathrm{TAP}_{N,n}(m,\varepsilon,\delta) \text{ is increasing in } \varepsilon \text{ and } \delta \text{ and sub-additive in } n,$$

the upper limit in (4.1) is increasing in ε and δ and decreasing in n, so the infimum over ε, δ, n can be replaced by the limit $\varepsilon, \delta \downarrow 0$ and $n \uparrow \infty$. Also, we can always choose $\varepsilon = \varepsilon_N$ and $\delta = \delta_N$ going to zero and $n = n_N$ going to infinity slowly enough so that

$$\inf_{\varepsilon,\delta,n} \limsup_{N\to\infty} \mathbb{E}\,\mathrm{TAP}_{N,n}(m,\varepsilon,\delta) = \limsup_{N\to\infty} \mathbb{E}\,\mathrm{TAP}_{N,n_N}(m,\varepsilon_N,\delta_N). \tag{9.1}$$

Using this representation and the equivalence $\mathrm{TAP}(\mu) = \mathbf{T}(\mu)$ in Theorem 3.4, to prove the lemma we need to show that

$$\limsup_{N\to\infty} \mathbb{E}\,\mathrm{TAP}_{N,n_N}(m,\varepsilon_N,\delta_N) \le \mathbf{T}(\mu). \tag{9.2}$$

First, consider the case $q = \int x^2\,\mu(dx) > 0$. Using the approximation in Lemma 6.2 and (6.5), it is enough to prove that

$$\limsup_{N\to\infty} \mathbb{E}\,F_{N,n_N}(m,\varepsilon_N,\delta_N) \le \mathbf{T}(\mu).$$

For any $v \in C([0,1])$, let $h_i = v(m_i)$ be the corresponding external field. Since $\|h\|/\sqrt{N} \le \|v\|_\infty < \infty$, by (3.14), deterministically,

$$\lim_{N\to\infty} \left|F_{N,n_N}(m,\varepsilon_N,\delta_N) - F^h_{N,n_N}(m,\varepsilon_N,\delta_N)\right| = 0. \tag{9.3}$$

Using that $F^h_{N,n_N}(m,\varepsilon_N,\delta_N) \le F^h_{N,1}(m,\varepsilon_N)$, by Theorem 3.2,

$$\limsup_{N\to\infty} \mathbb{E}\,F_{N,n_N}(m,\varepsilon_N,\delta_N) \le \limsup_{N\to\infty} \mathbb{E}\,F^h_{N,1}(m,\varepsilon_N) \le \mathcal{P}^v_\mu.$$

Taking infimum over $v \in V$ yields the assertion.

In the case when $q = \int x^2\,\mu(dx) = 0$, we can use the approximation in (6.9) and, in this case, Theorem 6.3 implies the claim without the need to introduce any external field, because the Parisi formula in Theorem 6.3 equals $\mathbf{T}(\delta_0)$. □

9.2 Proof of Lemma 4.2

Similarly to (9.2), to prove the lemma we need to show that

$$\liminf_{N\to\infty} \mathbb{E}\,\mathrm{TAP}_{N,n_N}(m,\varepsilon_N,\delta_N) \ge \mathbf{T}(\mu),$$

where ε_N, δ_N go to zero and n_N goes to infinity slowly enough.

Again, first, consider the case $q = \int x^2\,\mu(dx) > 0$. Using the approximation in Lemma 6.2 and (6.5), it is enough to prove that

$$\liminf_{N\to\infty} \mathbb{E}\,F_{N,n_N}(m,\varepsilon_N,\delta_N) \ge \mathbf{T}(\mu).$$

Consider the external field h defined through

$$h_i = v_{\zeta_0}(m_i) = \Psi\big(m_i,\zeta_0\big) \tag{9.4}$$

where ζ_0 is the minimizer found in Theorem 3.4. Since v is bounded on $[0, 1-\eta]$ and all $m_i \in [0, 1-\eta]$, by (3.14),

$$\liminf_{N\to\infty} \mathbb{E} F_{N,n_N}(m, \varepsilon_N, \delta_N) = \liminf_{N\to\infty} \mathbb{E} F^h_{N,n_N}(m, \varepsilon_N, \delta_N).$$

The Parisi formula in Theorem 3.2 implies that

$$\lim_{N\to\infty} \mathbb{E} F^h_{N,1}(m, \varepsilon_N, \delta_N) = \inf_{\lambda,\zeta} \mathcal{P}^{v_{\zeta_0}}_\mu(\lambda, \zeta) = \inf_{\lambda,\zeta} \overline{\mathcal{P}}^{v_{\zeta_0}}_\mu(\lambda, \zeta),$$

where the second equality holds because $\mathrm{supp}(\mu) \subseteq [0, 1-\eta]$ (so the functionals $\mathcal{P}$ and $\overline{\mathcal{P}}$ coincide). By our choice of ζ_0, Theorem 3.4 (iv) implies that the right-hand side equals $\overline{\mathcal{P}}^{v_{\zeta_0}}_\mu$, which equals to $\mathbf{T}(\mu)$ by Theorem 3.4(i). Furthermore, since zero is in the support of ζ_0 (Theorem 3.4(v)), one can argue that, for any fixed $n \geq 1$,

$$\lim_{N\to\infty} \mathbb{E} F^h_{N,n}(m, \varepsilon_N, \delta_N) = \lim_{N\to\infty} \mathbb{E} F^h_{N,1}(m, \varepsilon_N, \delta_N), \tag{9.5}$$

which, obviously, will finish the proof. This follows from a standard approximation argument by generic models, exactly as in [59], but, before we sketch it, let us notice that we are in the situation when

$$0 < q = \int x^2 \, \mu(dx) < 1 \quad \text{and} \quad \int x(1-x) \, \mu(dx) > 0,$$

which implies that:

(1) $(\lambda, \zeta) \to \mathcal{P}^{v_{\zeta_0}}_\mu(\lambda, \zeta)$ is strictly convex (see, e.g., (13.7) below);
(2) by equations (7.71) and (7.73) in Lemma 7.11, $\inf_{\lambda,\zeta} \mathcal{P}^{v_{\zeta_0}}_\mu(\lambda, \zeta)$ is achieved on λ that is uniformly bounded, $|\lambda| \leq L$.

For the specific model ξ_q we are considering above, v_{ζ_0} was chosen in an optimal way, so that the minimizer is $(0, \zeta_0)$. However, we will now vary the model ξ_q while keeping v_{ζ_0}, so the two items above refer to this case. In particular, by continuity and compactness, these items imply that the minimizer (let us denote it by $(\lambda_{\xi_q}, \zeta_{\xi_q})$) is unique and depends continuously on the model ξ_q (see e.g. [48, Corollary 4.2]). The arguments in [44, Section 3.7] require no modifications to show that, for generic models on the narrow band, the distribution of the overlap converges to some $\zeta^* \in \mathcal{M}_{0,1-q}$ and the limit of the free energy, via the Aizenman-Sims-Starr cavity computation in Section 7 above, is given by $\inf_\lambda \mathcal{P}^{v_{\zeta_0}}_\mu(\lambda, \zeta^*)$. On the other hand, by the Parisi formula in Theorem 3.2, this limit equals

$$\inf_{\lambda,\zeta} \mathcal{P}^{v_{\zeta_0}}_\mu(\lambda, \zeta) = \mathcal{P}^{v_{\zeta_0}}_\mu(\lambda_{\xi_q}, \zeta_{\xi_q}).$$

By uniqueness of the minimizer, $\zeta^* = \zeta_{\xi_q}$. Moreover, since this is the limiting distribution of the overlap, there can be no free energy cost of constraining the overlaps between n replicas to some fixed value in the support of ζ_{ξ_q}. On the other hand, in our model above, the external field v_{ζ_0} was chosen in such a way that zero is in the support of the minimizer ζ_0, so, when we approximate this model by generic models, by continuity of ζ_{ξ_q} in the model ξ_q, these generic models will

have points in the support very close to zero. As a result, for our model above, there can be no free energy cost of constraining the overlaps to be near zero and (9.5) must hold. For more details, see e.g. [48, Lemma 4.8].

In the case when $q = \int x^2 \mu(dx) = 0$, again, we can use the approximation in (6.9) and Theorem 6.3. The argument here is exactly the same, except we do not need to introduce the external field, because zero is already in the support of the Parisi measure of the original model without external field. □

10 TAP States Are Ancestral

In this section we will prove Theorem 2.4, which will follow from the following zero-temperature formula from [19], which is a generalization to soft spins of the zero-temperature result Jagannath-Sen [31, theorem 1.2] for discrete spins, which itself was derived from the positive temperature formulas with general prior spin distributions [42, 49, 50].

Define a functional $\mathcal{P}_q$ on $\mathbb{R} \times \mathcal{M}_{0,q}$ by

$$\mathcal{P}_q(\lambda, \gamma) = \Phi_\gamma^\lambda(0,0) - \frac{1}{2}\int_0^q s\xi''(s)\gamma(s)ds, \tag{10.1}$$

where, for a given λ, $\Phi_\gamma^\lambda(s,x)$ is defined as the solution of

$$\partial_s \Phi_\gamma^\lambda = -\frac{\xi''(s)}{2}\Big(\partial_{xx}\Phi_\gamma^\lambda + \gamma(s)\big(\partial_x \Phi_\gamma^\lambda\big)^2\Big) \tag{10.2}$$

on $[0,q) \times \mathbb{R}$, with the boundary condition

$$\Phi_\gamma^\lambda(q,x) := \max_{a\in[-1,1]}\big(ax + \lambda(a^2 - q) + \Lambda_{\zeta_*}(q,a)\big), \tag{10.3}$$

where $\Lambda_{\zeta_*}(q,a)$ was defined in (2.21) (recall that it is bounded and continuous on $[-1,1]$). Then [19, theorem 5] implies the following.

THEOREM 10.1. *For any $q \in [0,1]$, we have that*

(10.4)

$$\lim_{N\to\infty} \mathbb{E} \max_{\frac{1}{N}\|m\|^2 = q} \frac{1}{N}\Big(H_N(m) + \sum_{i=1}^N \Lambda_{\zeta_*}(q, m_i)\Big) = \inf_{(\lambda,\gamma)\in\mathbb{R}\times\mathcal{M}_{0,q}} \mathcal{P}_q(\lambda,\gamma).$$

PROOF OF THEOREM 2.4. Using this result, (2.32), Theorem 2.2, and Gaussian concentration, in order to prove Theorem 2.4, it is enough to prove that

$$\inf_{(\lambda,\gamma)\in\mathbb{R}\times\mathcal{M}_{0,q}} \mathcal{P}_q(\lambda,\gamma) - \frac{1}{2}\int_q^1 s\xi''(s)\zeta_*(s)ds \le \mathcal{P}(\zeta_*). \tag{10.5}$$

We will take $\gamma = \zeta_*$ (restricted to $[0,q]$) and take $\lambda = 0$. Then the function in the boundary condition (10.3) is $ax + \Lambda_{\zeta_*}(q,a)$ and, since $\Phi_{\zeta_*}(q,x)$ is convex, the definition (2.21) implies by conjugation that

$$\max_{a\in[-1,1]}\big(ax + \Lambda_{\zeta_*}(q,a)\big) = \Phi_{\zeta_*}(q,x).$$

Since the PDE in (10.2) with $\gamma = \zeta_*$ is the Parisi PDE for the original model, we get that

$$\mathcal{P}_q(0, \zeta_*) = \Phi_{\zeta_*}(0,0) - \frac{1}{2}\int_0^q s\xi''(s)\zeta_*(s)ds.$$

Finally,

$$\mathcal{P}_q(0,\zeta_*) - \frac{1}{2}\int_q^1 s\xi''(s)\zeta_*(s)ds = \Phi_{\zeta_*}(0,0) - \frac{1}{2}\int_0^1 s\xi''(s)\zeta_*(s)ds = \mathcal{P}(\zeta_*),$$

which finishes the proof. □

11 Generalized TAP Equations

In this section, we will prove the formula for the gradient of $\mathrm{TAP}(\mu_m)$ in Theorem 2.5 for all $m \in (-1,1)^N$. For $q \in [0,1]$, and $h \in \mathbb{R}^N$, define

$$\mathrm{TAP}(m,q,\zeta,h) := \frac{1}{N}\sum_{i=1}^N \big(\Phi_\zeta(q,h_i) - m_i h_i\big) - \frac{1}{2}\int_q^1 s\xi''(s)\zeta(s)ds. \tag{11.1}$$

If $\frac{1}{N}\|m\|^2 = q$ then the definition of $\mathrm{TAP}(\mu)$ in (2.25) implies that

$$\mathrm{TAP}(m) := \mathrm{TAP}(\mu_m) = \inf_{\zeta\in\mathcal{M}_{0,1},\, h\in\mathbb{R}^N} \mathrm{TAP}(m,q,\zeta,h). \tag{11.2}$$

If $m \in (-1,1)^N$, it is clear that we minimize over h in some cube $[-L,L]^N$, where L depends only on the largest value of $|m_i|$ (we will need this for compactness argument below). It is a standard fact that the functional $\mathrm{TAP}(m,q,\zeta,h)$ is strictly convex in h and convex in ζ. The reason it is not strictly convex in ζ is because the functional depends only on the restriction of ζ to $[q,1]$. For this reason, let us make a convention that, for a given q, we minimize over ζ fixed to be $\zeta(s) = 0$ for $s \in [0,q)$. Then the minimizer of the above functional for $m \in (-1,1)^N$ is unique and will be denoted by ζ_m, h_m, so that, for $q = \frac{1}{N}\|m\|^2$,

$$\mathrm{TAP}(m) = \mathrm{TAP}(m,q,\zeta_m,h_m). \tag{11.3}$$

If m converges to $m_0 \in (-1,1)^N$ then, by the continuity of $\mathrm{TAP}(\mu)$ proved in Theorem 3.4,

$$\lim_{m\to m_0} \mathrm{TAP}(m,q,\zeta_m,h_m) = \mathrm{TAP}(m_0,q_0,\zeta_{m_0},h_{m_0}).$$

Since $\mu_m \to \mu_{m_0}$, $q = \frac{1}{N}\|m\|^2 \to q_0 = \frac{1}{N}\|m_0\|^2$, and any subsequential limit of ζ_m is equal to zero on $[0,q_0)$ (by our convention above), the uniqueness of the minimizer implies that $\zeta_m \to \zeta_{m_0}$ and $h_m \to h_{m_0}$. With this observation, in order to compute the gradient of $\mathrm{TAP}(m)$, we will need two lemmas.

LEMMA 11.1. *Consider a metric space D and a function $f\colon(-\varepsilon,\varepsilon)\times D \to \mathbb{R}$. Suppose that there exists a function $d\colon(-\varepsilon,\varepsilon)\to D$ such that*

$$f(t,d(t)) = \inf_{d\in D} f(t,d), \tag{11.4}$$

and suppose that $d(t)$ is continuous at $t = 0$. Also, suppose that the right derivative $\partial_t^+ f(t, d)$ exists and is continuous at $(0, d(0))$. Then

$$\frac{d^+}{dt} f(t, d(t))\Big|_{t=0} = \partial_t^+ f(0, d(0)). \tag{11.5}$$

The same statement holds for left derivatives.

PROOF. Using (11.4), for $\alpha > 0$,

$$\begin{aligned}\frac{f(\alpha, d(\alpha)) - f(0, d(\alpha))}{\alpha} &\le \frac{f(\alpha, d(\alpha)) - f(0, d(0))}{\alpha} \\ &\le \frac{f(\alpha, d(0)) - f(0, d(0))}{\alpha}\end{aligned}$$

and, therefore,

$$\min_{s\in[0,\alpha]} \partial_t^+ f(s, d(\alpha)) \le \frac{f(\alpha, d(\alpha)) - f(0, d(0))}{\alpha} \le \max_{s\in[0,\alpha]} \partial_t^+ f(s, d(0)).$$

Letting $\alpha \downarrow 0$ finishes the proof. □

As we discussed above, TAP(m) is obtained by taking infimum over all $\zeta \in \mathcal{M}_{0,1}$ with $\zeta(s) = 0$ on $[0, q)$ and $h \in \mathbb{R}^N$. For any such ζ and $x \in \mathbb{R}$, define a stochastic process $u_{\zeta,x}$ on $[q, 1]$ by

$$u_{\zeta,x}(s) = \partial_x \Phi_\zeta(s, X_{\zeta,x}(s)),$$

where $(X_{\zeta,x}(s))_{q\le s\le 1}$ is the (strong) solution of the SDE

$$dX_{\zeta,x}(s) = \zeta(s)\xi''(s)\partial_x \Phi_\zeta(s, X_{\zeta,x}(s))ds + \xi''(s)^{1/2} dW_s$$

with the initial condition $X_{\zeta,x}(q) = x$.

LEMMA 11.2. *We have that*

$$\partial_{xx} \Phi_\zeta(q, x) = 1 - \int_{[q,1]} \mathbb{E} u_{\zeta,x}(l)^2 \, d\zeta(l). \tag{11.6}$$

PROOF. Recall from [4, Lemma 2] that there are two useful identities associated with the process $X_{\zeta,x}$, namely, for any $q \le s \le s' \le 1$,

$$\mathbb{E} u_{\zeta,x}^2(s') - \mathbb{E} u_{\zeta,x}^2(s) = \int_s^{s'} \xi''(r) \mathbb{E}\big(\partial_{xx} \Phi_\zeta(r, X_{\zeta,x}(r))\big)^2 \, dr, \tag{11.7}$$

and

$$\begin{aligned}&\partial_{xx} \Phi_\zeta(s', X_{\zeta,x}(s')) - \partial_{xx} \Phi_\zeta(s, X_{\zeta,x}(s)) \\ &= -\int_s^{s'} \xi''(r)\zeta(r)\big(\partial_{xx} \Phi_\zeta(r, X_{\zeta,x}(r))\big)^2 \, dr \\ &\quad + \int_s^{s'} \xi''(r)^{1/2} \partial_{xxx} \Phi_\zeta(r, X_{\zeta,x}(r)) dW_r.\end{aligned} \tag{11.8}$$

Using these and the Fubini theorem,

$$\begin{aligned}
&\mathbb{E}\partial_{xx}\Phi_\zeta(1, X_{\zeta,x}(1)) - \partial_{xx}\Phi_\zeta(q,x)\\
\text{(by (11.8))}\quad &= -\mathbb{E}\int_q^1 \xi''(r)\zeta(r)\big(\partial_{xx}\Phi_\zeta(r, X_{\zeta,x}(r))\big)^2\,dr\\
&= -\mathbb{E}\int_q^1 \xi''(r)\big(\partial_{xx}\Phi_\zeta(r, X_{\zeta,x}(r))\big)^2\Big(\int_{[q,r]}\zeta(dl)\Big)dr\\
&= -\mathbb{E}\int_{[q,1]}\int_l^1 \xi''(r)\big(\partial_{xx}\Phi_\zeta(r, X_{\zeta,x}(r))\big)^2\,dr\,d\zeta(l)\\
\text{(by (11.7))}\quad &= -\int_{[q,1]}\mathbb{E}\big(\partial_x\Phi_\zeta(1, X_{\zeta,x}(1))\big)^2 - \mathbb{E}\big(\partial_x\Phi_\zeta(l, X_{\zeta,x}(l))\big)^2\,d\zeta(l)\\
&= -\mathbb{E}\big(\partial_x\Phi_\zeta(1, X_{\zeta,x}(1))\big)^2 + \int_{[q,1]}\mathbb{E}\big(\partial_x\Phi_\zeta(l, X_{\zeta,x}(l))\big)^2\,d\zeta(l).
\end{aligned}$$

Noting that $\partial_{xx}\Phi_\zeta(1,x) = 1-\tanh^2(x) = 1-(\partial_x\Phi_\zeta(1,x))^2$, we get

$$\mathbb{E}\partial_{xx}\Phi_\zeta(1, X_{\zeta,x}(1)) = 1 - \mathbb{E}\big(\partial_x\Phi_\zeta(1, X_{\zeta,x}(1))\big)^2.$$

Combining these together completes our proof. □

PROOF OF THEOREM 2.5. Take $m_0 \in (-1,1)^N$ and consider the path $m_t = m_0 + tv$ for some $v \in \mathbb{R}^N$, which lies in $(-1,1)^N$ for $t \in (-\varepsilon,\varepsilon)$ for some small $\varepsilon > 0$, and denote $q_t = \frac{1}{N}\|m_t\|^2$. Let $D = \mathcal{M}_{0,1}\times\mathbb{R}^N$ and, with $d = (\zeta,h)$, let $f(t,d) := \mathrm{TAP}(m_t, q_t, \zeta, h)$. Let $d(t) := (\zeta_t, h_t) = (\zeta_{m_t}, h_{m_t})$ be the minimizers defined above, so that $\mathrm{TAP}(m_t) = f(t,d(t))$. Since $\nabla\,\mathrm{TAP}(m_0)\cdot v$ is the derivative of $f(t,d(t))$ at $t=0$, we can apply the above lemma once its assumptions are verified. The continuity of $d(t)$ at $t=0$ follows from the discussion above. To compute $\partial_t^+\,\mathrm{TAP}(m_t,q_t,\zeta,h)$, we need to compute the partial derivatives of $\mathrm{TAP}(m,q,\zeta,h)$ with respect to all m_i and q, which are the only parameters that depend on t for fixed (ζ,h). We will take the derivatives of m_t, q_t in t only at the end, using that, for $t=0$, we have $(m,q,\zeta,h) = (m_0,q_0,\zeta_0,h_0)$. First, right and left derivatives in q are equal to

$$\partial_q^\pm\,\mathrm{TAP}(m,q,\zeta,h) = \frac{1}{N}\sum_{i=1}^N \partial_q^\pm\Phi_\zeta(q,h_i) + \frac{1}{2}q\xi''(q)\zeta(q\pm 0),$$

where $\zeta(q\pm 0)$ are the one-sided limits of $\zeta(q)$. (Here, we consider both derivatives, because $q_t = \frac{1}{N}\|m_t\|^2$ may be increasing or decreasing with t.) Using the

Parisi PDE (2.6) for the first term, we can rewrite

$$\partial_q^{\pm} \mathrm{TAP}(m,q,\zeta,h) = -\frac{1}{N}\sum_{i=1}^{N}\frac{\xi''(q)}{2}\partial_{xx}\Phi_\zeta(q,h_i) + \frac{1}{2}\xi''(q)\zeta(q\pm 0)\Big(q - \frac{1}{N}\sum_{i=1}^{N}\big(\partial_x\Phi_\zeta(q,h_i)\big)^2\Big).$$

The only possible discontinuity on the right hand side is in the c.d.f. $\zeta(q)$. However, at $t=0$,

$$\partial_x\Phi_\zeta(q,h_i)\big|_{t=0} = \partial_x\Phi_{\zeta_0}(q_0,h_{0,i}) = m_{0,i}, \tag{11.9}$$

because $h_{0,i}$ is the minimizer and critical point of $\Phi_{\zeta_0}(q_0,h) - m_{0,i}h$ and, therefore,

$$\frac{1}{N}\sum_{i=1}^{N}\big(\partial_x\Phi_{\zeta_0}(q_0,h_{0,i})\big)^2 = \frac{1}{N}\sum_{i=1}^{N} m_{0,i}^2 = q_0.$$

This means that the one-sides derivatives above are continuous at (m_0,q_0,ζ_0,h_0) and

$$\partial_q^{\pm}\mathrm{TAP}(m,q,\zeta,h)\big|_{t=0} = -\frac{1}{N}\sum_{i=1}^{N}\frac{\xi''(q_0)}{2}\partial_{xx}\Phi_{\zeta_0}\big(q_0,\overline{\Psi}(q_0,m_{0,i},\zeta_0)\big), \tag{11.10}$$

since the minimizer $h_{0,i} = \overline{\Psi}(q_0,m_{0,i},\zeta_0)$ was defined in (2.22). The derivative $\mathrm{TAP}(m,q,\zeta,h)$ in m_i equals $-h_i$, which is continuous and

$$\partial_{m_i}\mathrm{TAP}(m,q,\zeta,h)\big|_{t=0} = -\frac{1}{N}h_{0,i} = -\frac{1}{N}\overline{\Psi}(q_0,m_{0,i},\zeta_0). \tag{11.11}$$

If we denote

$$C(m_0) := \xi''(q_0)\frac{1}{N}\sum_{i=1}^{N}\partial_{xx}\Phi_{\zeta_0}\big(q_0,\overline{\Psi}(q_0,m_{0,i},\zeta_0)\big) = \xi''(q_0)\frac{1}{N}\sum_{i=1}^{N}\partial_{xx}\Phi_{\zeta_0}\big(q_0,h_{0,i}\big)$$

and use that $\frac{d}{dt}q|_{t=0} = 2m_0\cdot v$, combining the above we have that

$$\begin{aligned}\nabla\,&\mathrm{TAP}(m_0)\cdot v \\ &= \frac{d}{dt}q\cdot\partial_q\,\mathrm{TAP}(m,q,\zeta,h)\big|_{t=0} + \sum_{i=1}^{N}\frac{d}{dt}m_i\cdot\partial_{m_i}\,\mathrm{TAP}(m,q,\zeta,h)\big|_{t=0} \\ &= -\frac{1}{N}\big(\overline{\Psi}(q_0,m_{0,i},\zeta_0) + C(m_0)m_{0,i}\big)_{i\le N}\cdot v\end{aligned}$$

for all $v \in \mathbb{R}^N$. To complete the proof, it remains to verify that

$$C(m_0) = \xi''(q_0)\int_{q_0}^1 \zeta_0(s)ds. \tag{11.12}$$

For a fixed $m_0 \in (-1,1)^N$ with $q_0 = \frac{1}{N}\|m_0\|^2$, let us find the minimizers ζ_0 and h_0 of the strictly convex variational problem (11.2), with the above convention that we optimize over ζ fixed to be $\zeta(s) = 0$ for $s \in [0, q)$. Then the directional derivative with respect to ζ at ζ_0 in the direction of any other distributions must be non-negative. If, for a given $\zeta \in \mathcal{M}_{0,1}$, we consider the path $\zeta_b = \zeta_0 + b(\zeta - \zeta_0)$ parametrized by $b \in [0,1]$, one can compute the directional derivative in a standard way (see [16, theorem 2] or [48, lemma 4.11]) to get

$$\begin{aligned}&\frac{d}{db}\,\mathrm{TAP}(m_0, q_0, \zeta_b, h_0)\Big|_{b=0^+}\\ &\quad= \frac{1}{2}\int_q^1 \xi''(s)\Big(\frac{1}{N}\sum_{i=1}^N \mathbb{E}u_{\zeta_0,h_{0,i}}(s)^2 - s\Big)\big(\zeta(s) - \zeta_0(s)\big)ds \geq 0.\end{aligned}$$

Now, if we vary over all possible ζ, the minimality of ζ_0 implies that whenever $s \in [q,1]$ is in the support of ζ_0, we must have that (see, e.g., [16, 33])

$$\frac{1}{N}\sum_{i=1}^N \mathbb{E}u_{\zeta_0,h_{0,i}}(s)^2 = s.$$

Plugging this into (11.6) with $\zeta = \zeta_0$ and initial condition $X_{\zeta,x}(q) = x = h_{0,i}$ and averaging over i,

$$\frac{1}{N}\sum_{i=1}^N \partial_{xx}\Phi_{\zeta_0}(q, h_{0,i}) = 1 - \int_{[q,1]} l\, d\zeta(l).$$

Using that

$$\begin{aligned}\int_{[q,1]} l\, d\zeta(l) &= \int_{[q,1]}\Big(\int_q^l dw\Big)d\zeta(l) + q = \int_q^1\int_{[w,1]} d\zeta(l)dw + q\\ &= \int_q^1 (1 - \zeta(w^-))dw + q = 1 - \int_q^1 \zeta(w)dw\end{aligned}$$

finishes the proof of (11.12). □

12 Classical TAP Correction

We proceed to establish Propositions 3.5 and 3.6 and Corollary 2.3.

PROOF OF PROPOSITION 3.5. Assume that $(0, \delta_0)$ is the minimizer to $\overline{\mathcal{P}}_\mu^{v_{\mathrm{RS}}}$. Let $\zeta \in \mathcal{M}_{0,1-q}$ be fixed. For $b \in [0,1]$, define

$$\zeta_b = (1-b)\delta_0 + b\zeta.$$

Then from the minimality of δ_0, one gets that (see Remark 13.7)

$$\frac{d}{db}\overline{\mathcal{P}}_\mu^{\upsilon_{\mathrm{RS}}}(0,\zeta_b)\Big|_{b=0^+} = \frac{1}{2}\int_0^{1-q} \xi_q''(s)(\gamma_\mu(s)-s)(\zeta(s)-\delta_0(s))ds \geq 0.$$

Next, by Fubini's theorem, write

$$\begin{aligned}
0 &\leq \int_0^{1-q} \xi_q''(s)(\gamma_\mu(s)-s)(\zeta(s)-\delta_0(s))ds \\
&= \int_0^{1-q} \xi_q''(s)(\gamma_\mu(s)-s)\left(\int_0^s (\zeta(dr)-\delta_0(dr))\right)ds \\
&= \int_0^{1-q}\int_r^{1-q} \xi_q''(s)(\gamma_\mu(s)-s)ds(\zeta(dr)-\delta_0(dr)) \\
&= \int_0^{1-q}\int_r^{1-q} \xi_q''(s)(\gamma_\mu(s)-s)ds\zeta(dr) \\
&\quad - \int_0^{1-q} \xi_q''(s)(\gamma_\mu(s)-s)ds \\
&= -\int_0^{1-q} \Gamma_\mu(r)\zeta(dr).
\end{aligned} \tag{12.1}$$

Since this inequality holds for all $\zeta \in \mathcal{M}_{0,1-q}$, it follows that $\Gamma_\mu(r) \leq 0$ for all $r \in [0, 1-q]$. Conversely, if $\Gamma_\mu(r) \leq 0$ for all $0 \leq r \leq 1-q$, one can reverse the argument to get that

$$\frac{d}{db}\overline{\mathcal{P}}_\mu^{\upsilon_{\mathrm{RS}}}(0,\zeta_b)\Big|_{b=0^+} \geq 0$$

for all $\zeta \in \mathcal{M}_{0,1-q}$. Since $\overline{\mathcal{P}}_\mu^{\upsilon_{\mathrm{RS}}}(0,\cdot)$ is a strictly convex functional (using the assumption $\mu \neq \delta_1$), this implies that δ_0 is the unique minimizer of $\overline{\mathcal{P}}_\mu^{\upsilon_{\mathrm{RS}}}(0,\cdot)$. In order to show that $(\lambda_\mu^{\upsilon_{\mathrm{RS}}}, \zeta_\mu^{\upsilon_{\mathrm{RS}}}) = (0,\delta_0)$ we split our discussion into two cases: the support of μ contains a point in $(0,1)$ and μ is supported only on $\{0,1\}$ with $\mu(\{0\}) > 0$. In the latter case, recall that we defined $\lambda_\mu^{\upsilon_{\mathrm{RS}}} = 0$ and $\zeta_\mu^{\upsilon_{\mathrm{RS}}}$ to be the (unique) minimizer of $\overline{\mathcal{P}}_\mu^{\upsilon_{\mathrm{RS}}}(0,\cdot)$. Hence, $(\lambda_\mu^{\upsilon_{\mathrm{RS}}}, \zeta_\mu^{\upsilon_{\mathrm{RS}}}) = (0,\delta_0)$. In the former case, Remark 13.5 shows that there exists a unique minimizer $(\lambda_\mu^{\upsilon_{\mathrm{RS}}}, \zeta_\mu^{\upsilon_{\mathrm{RS}}})$ of $\overline{\mathcal{P}}_\mu^{\upsilon_{\mathrm{RS}}}(\cdot,\cdot)$. Since by definition υ_{RS} satisfies $\Phi_{a,\delta_0}(0,\upsilon_{\mathrm{RS}}(a)) = 0$ for all $a \in [0,1)$, this implies that $\partial_\lambda \mathcal{P}_\mu^{\upsilon_{\mathrm{RS}}}(0,\delta_0) = 0$. This together with the fact that δ_0 is the minimizer of $\overline{\mathcal{P}}_\mu^{\upsilon_{\mathrm{RS}}}(0,\cdot)$ implies that $(\lambda_\mu^{\upsilon_{\mathrm{RS}}}, \zeta_\mu^{\upsilon_{\mathrm{RS}}}) = (0,\delta_0)$. This establishes the equivalence conditions in the statement of Proposition 3.5.

Finally, we compute $\overline{\mathcal{P}}_\mu^{\upsilon_{\mathrm{RS}}}(0,\delta_0)$ assuming that either of the conditions in the equivalence holds. Recall the explicit expression of Φ_{a,δ_0} from (3.40) and the

particular choice of v_{RS} from (3.41). From these, it can be checked that

$$\begin{aligned}\Phi_{a,\delta_0}(0, v_{\mathrm{RS}}(a)) &= \frac{1-a^2}{2}\xi_q'(1-q) - a\tanh^{-1}(a) + \log 2\cosh(\tanh^{-1}(a)) \\ &= \frac{1-a^2}{2}\xi_q'(1-q) - I(a).\end{aligned}$$

From this,

$$\begin{aligned}&\int_{[0,1)} \Phi_{a,\delta_0}(0, v_{\mathrm{RS}}(a))d\mu(a) \\ &\quad= \frac{1}{2}\xi_q'(1-q)\int_{[0,1)}(1-a^2)d\,d\mu(a) - \int_{[0,1)} I(a)d\mu(a) \\ &\quad= \frac{1}{2}\xi_q'(1-q)\int(1-a^2)da - \int I(a)d\mu(a) \\ &\quad= \frac{1}{2}\xi_q'(1-q)(1-q) - \int I(a)d\mu(a),\end{aligned}$$

where we have used that $I(a)$ is continuous at 1 and $I(1) = 0$. In addition, using integration by parts

$$\int_0^{1-q} \xi_q''(s)s\,ds = \xi_q'(1-q)(1-q) - \xi_q(1-q).$$

Therefore,

$$\begin{aligned}&\overline{\mathcal{P}}_\mu^{v_{\mathrm{RS}}}(0, \delta_0) \\ &\quad= \int_{[0,1)} \Phi_{a,\delta_0}(0, v_{\mathrm{RS}}(a))d\mu(a) - \frac{1}{2}\int_0^{1-q} \xi_q''(s)s\,ds \\ &\quad= \frac{1}{2}\xi_q'(1-q)(1-q) - \int I(a)d\mu(a) - \frac{1}{2}\xi_q'(1-q)(1-q) + \frac{1}{2}\xi_q(1-q) \\ &\quad= -\int I(a)d\mu(a) + \frac{1}{2}\big(\xi(1) - \xi(q) - \xi'(q)(1-q)\big),\end{aligned}$$

where we used $\xi_q(1-q) = \xi(1) - \xi(q) - \xi'(q)(1-q)$. This establishes the desired formula for $\overline{\mathcal{P}}_\mu^{v_{\mathrm{RS}}}(0, \delta_0)$. Finally, since $\mu \neq \delta_1$ and $(0, \delta_0)$ is the minimizer to $\overline{\mathcal{P}}_\mu^{v_{\mathrm{RS}}}$, it follows from Theorem 3.4 that $\mathbf{T}(\mu) = \overline{\mathbf{T}}(\mu) = \overline{\mathcal{P}}_\mu^{v_{\mathrm{RS}}}$. □

PROOF OF COROLLARY 2.3. For $m \in [-1, 1]^N$ and $\mu \in \mathcal{M}_{0,1}$, denote $q_m = \|m\|_2^2/N$ and $q = \int a^2\,d\mu(a)$. Recall from [19] that if q_{EA} is the largest point in the support of the original Parisi measure of the Parisi formula for F_N, then almost surely

$$\begin{aligned}&\lim_{N\to\infty} F_N \\ &\quad= \lim_{\varepsilon\downarrow 0}\lim_{N\to\infty}\sup_{m\in[-1,1]^N:|q_m - q_{\mathrm{EA}}|\le\varepsilon}\left(\frac{H_N(m)}{N} - \int I(a)d\mu_m(a) + C(q_m)\right).\end{aligned} \tag{12.2}$$

Also recall from Theorem 2.2 that

$$\lim_{N\to\infty} F_N = \lim_{\varepsilon\downarrow 0}\lim_{N\to\infty} \sup_{m\in[-1,1]^N:|q_m-q_{\mathrm{EA}}|\le\varepsilon} \left(\frac{H_N(m)}{N} + \bar{\mathbf{T}}(\mu_m)\right). \tag{12.3}$$

Since

$$\bar{\mathbf{T}}(\mu_m) \le \bar{P}^{\upsilon_{\mathrm{RS}}}_{\mu_m}(0,\delta_0) = -\int I(a)d\mu_m(a) + C(q_m), \quad \forall m\in[-1,1]^N, \tag{12.4}$$

we see that

$$\lim_{\varepsilon\downarrow 0}\lim_{N\to\infty} \sup_{m\in[-1,1]^N:|q_m-q_{\mathrm{EA}}|\le\varepsilon} \left(-\int I(a)d\mu_m(a) + C(q_m) - \bar{\mathbf{T}}(\mu_m)\right) = 0. \tag{12.5}$$

We claim that for any $\varepsilon_N \downarrow 0$,

$$\limsup_{N\to\infty} \sup_{m\in[-1,1]^N:|q_m-q_{\mathrm{EA}}|\le\varepsilon_N} \sup_{0\le s\le 1-q_m} \Gamma_{\mu_m}(s) \le 0.$$

If this is not true, then there exists some $\delta > 0$ and a sequence m^N with $|q_{m^N} - q_{\mathrm{EA}}| \le \varepsilon_N$ such that μ_{m^N} converges to certain μ weakly and

$$\sup_{0\le s\le 1-q_{m^N}} \Gamma_{\mu_{m^N}}(s) \ge \delta.$$

From these, passing to the limit gives

$$\int a^2\, d\mu(a) = q_{\mathrm{EA}} \text{ and } \sup_{0\le s\le 1-q_{\mathrm{EA}}} \Gamma_\mu(s) \ge \delta.$$

Note that $q_{\mathrm{EA}} < 1$ implies $\mu \ne \delta_1$. From this, the above display, and Proposition 3.5, we arrive at

$$\bar{\mathbf{T}}(\mu) < \bar{\mathcal{P}}^{\upsilon_{\mathrm{RS}}}_{\mu}(0,\delta_0) = -\int I(a)d\mu(a) + C(q).$$

Since these two sides are uniformly continuous functions of μ, we see that

$$\liminf_{N\to\infty}\left(-\int I(a)d\mu_m(a) + C(q_{m^N}) - \bar{\mathbf{T}}(\mu_{m^N})\right) > 0,$$

which contradicts (12.5). This establish our claim.

Now from the above claim, for any $\eta > 0$, there exists an $\varepsilon > 0$ such that as long as N is large enough, if $m \in [-1,1]^N$ satisfies $|q_m - q_{\mathrm{EA}}| \le \varepsilon$, then

$$\sup_{0\le s\le 1-q_m} \Gamma_{\mu_m}(s) \le \eta.$$

From this, (12.3), and Theorem 2.2,

$$\lim_{N\to\infty} F_N = \lim_{\eta\downarrow 0}\lim_{N\to\infty} \sup_{\substack{m\in[-1,1]^N:\\ \sup_{0\le s\le 1-q_m}\Gamma_{\mu_m}(s)\le\eta}} \left(\frac{H_N(m)}{N} + \bar{\mathbf{T}}(\mu_{\mu_m})\right).$$

Recall (12.4). Our proof will be completed if it is established that

(12.6)
$$\lim_{\eta\downarrow 0}\lim_{N\to\infty}\sup_{\substack{m\in[-1,1]^N:\\ \sup_{0\le s\le 1-q_m}\Gamma_{\mu_m}(s)\le\eta}}\left(-\int I(a)d\mu_m(a)+C(q_m)-\bar{\mathbf{T}}(\mu_m)\right)=0.$$

The argument of proving this is essentially the same as the above claim. Assume on the contrary that there exist $\eta_N\downarrow 0$, $\delta>0$ and $m^N\in[-1,1]^N$ with

$$\sup_{0\le s\le 1-q_{m^N}}\Gamma_{\mu_{m^N}}(s)\le\eta_N$$

such that μ_{m^N} weakly converges to some μ and

$$-\int I(a)\mu_{m^N}(da)+C(q_{m^N})-\bar{\mathbf{T}}(\mu_{m^N})\ge\delta,\quad\forall N\ge 1.$$

From these, we see that by passing to the limit,

$$\sup_{0\le s\le 1-q}\Gamma_\mu(s)\le 0$$

and

$$-\int I(a)\mu(da)+C(q)-\bar{\mathbf{T}}(\mu)\ge\delta.$$

If $\mu\neq\delta_1$, then these contradict Proposition 3.5. If $\mu=\delta_1$, then $q=1$ and in this case, it can be clearly checked that

$$\bar{\mathbf{T}}(\mu)=0=-\int I(a)d\mu(a)+C(q),$$

which again contradict to the above inequality. Hence (12.6) must be valid. □

PROOF OF PROPOSITION 3.6. Note that $\xi_q(s)=\beta^2s^2/2$ and

$$\Gamma_\mu(s)=\beta^2\int_0^s(\gamma_\mu(s)-s)ds.$$

Assume that $(\lambda_\mu^{v_{\mathrm{RS}}},\zeta_\mu^{v_{\mathrm{RS}}})=(0,\delta_0)$. From Proposition 3.5, it can be seen that $\Gamma_\mu(s)$ attains the global maximum at 0. On the other hand, it can also be checked that $\frac{d}{ds}\Gamma_\mu(0)=\beta^2\gamma_\mu(0)=0$ by a direct computation. From these, it follows that the second derivative of Γ_μ at zero is not positive. Now, following the same computation as [3, Proposition 3], this second derivative can be computed as

$$\frac{d^2}{ds^2}\Gamma_\mu(0)=\beta^2\left(\beta^2\int\mathbb{E}\partial_{xx}\Phi_{a,\delta_0}(0,v_{\mathrm{RS}}(a))^2\,d\mu(a)-1\right)\le 0.$$

Consequently,

$$\beta^2\int\mathbb{E}\partial_{xx}\Phi_{a,\delta_0}(0,v_{\mathrm{RS}}(a))^2\,d\mu(a)\le 1.$$

Finally, since

$$\partial_{xx}\Phi_{a,\delta_0}(0, v_{\mathrm{RS}}(a)) = 1 - \tanh^2(v_{\mathrm{RS}}(a) - at^2(0)) = 1 - a^2, \forall a \in [0,1),$$

we arrive at the Plefka condition by plugging this equation into the above inequality. □

13 Analytical Results

This section is devoted to handling some basic properties of the effective field $\Psi(a,\zeta)$ and the Parisi functional $\overline{\mathcal{P}}^v_\mu(\lambda,\zeta)$.

13.1 Basic facts about $\Phi_{a,\zeta}$

Recall that the PDE solutions $\Phi_{a,\zeta}$ and Φ_ζ are connected by (8.4). In order to state several useful properties of $\Phi_{a,\zeta}$ in the present paper, we first recall some well-known properties of the PDE solution Φ_ζ. First of all, Φ_ζ satisfies $\sup_{\zeta\in\mathcal{M}_{0,1}} \|\partial_{x^k}\Phi_\zeta\|_\infty < \infty$ for all $k \geq 1,2,3$. Second, Φ_ζ can be written as a stochastic optimization problem. Third, for any $t \in [0,1]$, $(\zeta,x) \in \mathcal{M}_{0,1}\times\mathbb{R} \mapsto \partial_x^k\Phi_\zeta(t,x)$ for $k = 0,1,2,3$ is Lipschitz and $(\zeta,x) \in \mathcal{M}_{0,1}\times\mathbb{R} \mapsto \Phi_\zeta(t,x)$ is strictly convex. Lastly, the directional derivative of $\Phi_\zeta(t,x)$ in (ζ,x) exists and admits an explicit formula in terms of the optimal process appearing in the stochastic control representation of Φ_ζ. See [3,4,16,32] for these results. Due to the equation (8.4), it can be checked immediately that the following statements are valid for any $\mu \in \mathcal{M}_{0,1}$ with $q = \int a^2\,d\mu(a)$:

(I) Regularity: We have that

$$\sup_{(a,\zeta,t,x)\in[0,1]\times\mathcal{M}_{0,1-q}\times[1-q]\times\mathbb{R}} \Big\{ |\partial_x\Phi_{a,\zeta}(t,x)|, |\partial_{xx}\Phi_{a,\zeta}(t,x)|, |\partial_{xx}\Phi_{a,\zeta}(t,x)|, |\partial_{ax}\Phi_{a,\zeta}(t,x)| \Big\} < \infty. \tag{13.1}$$

(II) Stochastic optimal control: The quantity $\Phi_{a,\zeta}(0,x)$ can be expressed as a stochastic optimal control problem, which states that

$$\begin{aligned}\Phi_{a,\zeta}(0,x) &= \sup_u \Big(\mathbb{E} f\Big(a, x + \int_0^{1-q} \xi_q''(s)\zeta(s)u(s)ds + \int_0^{1-q} \xi_q''(s)^{1/2} dW_s\Big) \\ &\quad - \frac{1}{2}\int_0^{1-q} \xi_q''(s)\zeta(s)\mathbb{E}u(s)^2\,ds\Big),\end{aligned} \tag{13.2}$$

where the supremum is over all progressively measurable processes

$$u = (u(s))_{0\leq s\leq 1-q}$$

with respect to the standard Brownian motion $W = (W_s)_{0\le s\le 1-q}$ and with

$$\sup_{s\in[0,1-q]} |u(s)| \le 2.$$

Here the optimal process is attained by $u_{a,\zeta,x}(s) = \partial_x \Phi_{a,\zeta}(s, X_{a,\zeta,x}(s))$, where $X_{a,\zeta,x}$ is the solution to the following SDE with initial condition $X_{a,\zeta,x}(0) = x$,

$$\begin{aligned} dX_{a,\zeta,x}(s) &= \xi_q''(s)\zeta(s)\partial_x \Phi_{a,\zeta}(s, X_{a,\zeta,x}(s))ds \\ &\quad + \xi_q''(s)^{1/2} dW_s, \quad \forall 0 \le s \le 1-q. \end{aligned} \tag{13.3}$$

(III) Lipschitz property: For any $(m,n) \in \{0,1\}\times\{0,1,2\}$, $(a,\zeta,x) \in [0,1]\times \mathcal{M}_{0,1-q}\times\mathbb{R} \mapsto \partial_{a^m x^n}\Phi_{a,\zeta}(t,x)$ is Lipschitz in the sense that

$$\begin{aligned} &\left|\partial_{a^m x^n}\Phi_{a,\zeta}(0,x) - \partial_{a^m x^n}\Phi_{a',\zeta'}(0,x')\right| \\ &\quad \le C\left(|a-a'| + |x-x'| + \int_0^{1-q} |\zeta-\zeta'|ds\right) \end{aligned} \tag{13.4}$$

for some universal constant $C > 0$ depending only on ξ.

(IV) Convexity: For any $\zeta_0, \zeta_1 \in \mathcal{M}_{0,1-q}$ and $x_0, x_1 \in \mathbb{R}$, define

$$\begin{aligned} \zeta_b &= (1-b)\zeta_0 + b\zeta_1, \\ x_b &= (1-b)x_0 + bx_1, \end{aligned} \tag{13.5}$$

for $b \in [-1,1]$. For any $a \in [0,1]$, we have that

$$\Phi_{a,\zeta_b}(0,x_b) \le (1-b)\Phi_{a,\zeta_0}(0,x_0) + b\Phi_{a,\zeta_1}(0,x_1). \tag{13.6}$$

Furthermore, whenever $(\zeta_0,x_0) \ne (\zeta_1,x_1)$ and $b \in (0,1)$, this inequality is strict,

$$\Phi_{a,\zeta_b}(0,x_b) < (1-b)\Phi_{a,\zeta_0}(0,x_0) + b\Phi_{a,\zeta_1}(0,x_1). \tag{13.7}$$

(V) Directional derivative: Recall the convex combination in (13.5). The derivative of $\Phi_{a,\cdot}(0,\cdot)$ exists and is equal to

$$\begin{aligned} \frac{d}{db}\Phi_{a,\zeta_b}(0,x_b) &= \frac{1}{2}\int_0^{1-q} \xi_q''(s)(\zeta_1(s)-\zeta_0(s))\mathbb{E}u_{a,\zeta_b,x_b}(s)^2\, ds \\ &\quad + (x_1-x_0)\partial_x\Phi_{a,\zeta_b}(0,x_b) \end{aligned} \tag{13.8}$$

for any $a \in [0,1]$ and $b \in (0,1)$, where u_{a,ζ_b,x_b} is the optimal process of (13.2) with $x = x_b$ and $\zeta = \zeta_b$. Furthermore, the right derivative of $\Phi_{a,\zeta_b}(0,x_b)$ also exists at $b = 0$ and it is equal to the above formula with $b = 0$.

13.2 Basic properties of $\Psi(a,\zeta)$

Now we use the properties listed in the above subsection to study a number of key features of $\Psi(a,\zeta)$ defined in (3.26).

LEMMA 13.1. *For any $q \in [0,1]$, Ψ is continuous on $[0,1) \times \mathcal{M}_{0,1-q}$.*

PROOF. Let $a_0 \in [0,1)$, $\zeta_0 \in \mathcal{M}_{0,1-q}$ be fixed. By definition of $\Psi(a_0,\zeta_0)$, for any $\varepsilon > 0$,

$$\delta := \min_{x \in I(\varepsilon)^c} |\partial_x \Phi_{a_0,\zeta_0}(0,x)| > 0,$$

where $I(\varepsilon)^c$ is the complement of an open interval defined by

$$I(\varepsilon) = (\Psi(a_0,\zeta_0) - \varepsilon, \Psi(a_0,\zeta_0) + \varepsilon).$$

From the Lipschitz property (13.4) with $(m,n) = (0,1)$, we see that if

$$|a - a_0| + \int_0^{1-q} |\zeta(s) - \zeta_0(s)| ds < \frac{\delta}{2C},$$

then for any $x \in I(\varepsilon)^c$, we have

$$\begin{aligned} &|\partial_x \Phi_{a,\zeta}(0,x)| \\ &\geq |\partial_x \Phi_{a_0,\zeta_0}(0,x)| - C\left(|a - a_0| + \int_0^{1-q} |\zeta(s) - \zeta_0(s)| ds\right) \geq \frac{\delta}{2}. \end{aligned}$$

This implies that

$$\min_{x \in I(\varepsilon)^c} |\partial_x \Phi_{a,\zeta}(0,x)| \geq \frac{\delta}{2}.$$

Since $\partial_x \Phi_{a,\zeta}(0, \Psi(a,\zeta)) = 0$, it follows that $\Psi(a,\zeta) \in I(\varepsilon)$. Hence, $\Psi(a,\zeta)$ is continuous on $[0,1) \times \mathcal{M}_{0,1-q}$. □

Recall the definition of the effective field $v_\zeta = \Psi(a,\zeta)$ from (3.31).

LEMMA 13.2. *Let $q \in [0,1]$. For any $\zeta \in \mathcal{M}_{0,1-q}$, v_ζ is a well-defined strictly increasing function with $v_\zeta(0) = 0$. In addition, for any $l \in (0,1)$,*

$$\sup_{(a,\zeta) \in [0,l] \times \mathcal{M}_{0,1-q}} \partial_a v_\zeta(a) < \infty.$$

PROOF. Let $q \in [0,1]$ be fixed. Recall from (8.4) that if $\zeta \in \mathcal{M}_{0,1-q}$ and $\zeta_0 \in \mathcal{M}_{0,1}$ satisfy $\zeta(t) = \zeta_0(q+t)$ for $t \in [0, 1-q]$, then for any $a \in [-1,1]$ and $x \in \mathbb{R}$,

$$\begin{aligned} \Phi_{a,\zeta}(0,x) = \Phi_\zeta\left(q, x - a\int_q^1 \xi''(s)\zeta_0(s)ds\right) - ax \\ + \frac{a^2}{2}\int_q^1 \xi''(s)\zeta_0(s)ds. \end{aligned} \tag{13.9}$$

Since $\lim_{|x|\to\infty} \Phi_{\zeta_0}(q,x) = \infty$, we see that $\lim_{|x|\to\infty} \Phi_{a,\zeta}(0,x) = \infty$ for all $a \in [0,1)$ and $\zeta \in \mathcal{M}_{0,1-q}$. On the other hand, we also know that $\Phi_{a,\zeta}(0,\cdot)$ is a

strictly convex function, by (13.7). These imply that for any $a \in [0,1)$, $\Phi_{a,\xi}(0,\cdot)$ has only one critical point, so v_ξ is well-defined. In particular, when $a = 0$, $\Phi_{a,\xi}(0,\cdot)$ is an even function so that $v_\xi(0) = 0$.

Next, we show that v_ξ is strictly increasing. First, note that $\partial_{xx}\Phi_{\zeta_0}(q,x) > 0$ for all $x \in \mathbb{R}$ since $\Phi_{\zeta_0}(q,\cdot)$ is strictly convex. From this and (13.9), for any $(a,\zeta,x) \in [0,1] \times \mathcal{M}_{0,1-q} \times \mathbb{R}$,

$$\begin{aligned}\partial_{ax}\Phi_{a,\xi}(0,x) &= -\left(\int_0^1 \xi''(s)\zeta_0(s)ds\right)\partial_{xx}\Phi_{\zeta_0} \\ &\quad \times \left(0, x - a\int_0^1 \xi''(s)\zeta_0(s)ds\right) < 0, \\ \partial_{xx}\Phi_{a,\xi}(0,x) &= \partial_{xx}\Phi_{\zeta_0}\left(0, x - a\int_0^1 \xi''(s)\zeta_0(s)ds\right) > 0.\end{aligned} \tag{13.10}$$

Consequently, a direct differentiation of $\partial_x\Phi_{a,\xi}(0, v_\xi(a)) = 0$ in $a \in [0,1)$ and using (13.10) yield that

$$\partial_a v_\xi(a) = -\frac{\partial_{ax}\Phi_{a,\xi}(0,x)}{\partial_{xx}\Phi_{a,\xi}(0,x)}\bigg|_{x=v_\xi(a)} > 0. \tag{13.11}$$

Hence, v_ξ is strictly increasing on $[0,1)$.

Finally, we prove the uniform upper bound for $\partial_a v_\xi(a)$. Fix $l \in (0,1)$. By Lemma 13.1, $v_\xi(a) \le L$ for all $a \le l$ and all ζ, for some large enough L. Therefore,

$$\begin{aligned}&\inf_{(a,\xi)\in[0,l]\times\mathcal{M}_{0,1-q}} \partial_{xx}\Phi_{a,\xi}(0, v_\xi(a)) \\ &\quad\ge \inf_{(a,\xi,x)\in[0,1]\times\mathcal{M}_{0,1-q}\times[0,L]} \partial_{xx}\Phi_{a,\xi}(0,x) > 0,\end{aligned}$$

by (13.10), continuity (13.4) and compactness. On the other hand, from (13.1),

$$\sup_{(a,\xi,x)\in[-1,1]\times\mathcal{M}_{0,1-q}\times\mathbb{R}} \left|\partial_{ax}\Phi_{a,\xi}(0,x)\right| < \infty.$$

From these inequalities and (13.11), $\sup_{(a,\xi)\in[0,l]\times\mathcal{M}_{0,1-q}} \partial_a v_\xi(a) < \infty$. □

LEMMA 13.3. *There exist positive constants $c_1, c_1' \in \mathbb{R}$ and $c_2, c_2' > 0$ such that*

$$c_1' + c_2' \tanh^{-1}(a) \le \Psi(a,\zeta) \le c_1 + c_2 \tanh^{-1}(a) \tag{13.12}$$

for all $a \in [0,1)$, $q \in [0,1]$ and $\zeta \in \mathcal{M}_{0,1-q}$.

PROOF. Note that $f(t,x) := \partial_x\Phi_{a,\xi}(t,x)$ satisfies the equation

$$\partial_t f = -\frac{\xi_q''(t)}{2}\big(\partial_{xx} f + 2\zeta(t)(\partial_x f)(\partial_x\Phi_{a,\xi})\big)$$

with $f(1-q,x) = -a + \tanh x$. Using the Feyman-Kac formula,

$$\partial_x\Phi_{a,\xi}(0,x) = -a + \mathbb{E}\tanh X_{a,\xi,x}(1-q), \tag{13.13}$$

where $X_{a,\xi,x}$ is defined through (13.3). Note that since $|u(s)| \le 2$ and $|\zeta| \le 1$, we have that $x + z_- \le X_{a,\xi,x}(1-q) \le x + z_+$, where

$$z_- = -2\xi_q'(1-q) + \int_0^{1-q} \xi_q''(s)^{1/2}\, dW_s,$$
$$z_+ = 2\xi_q'(1-q) + \int_0^{1-q} \xi_q''(s)^{1/2}\, dW_s.$$

Using that $\xi_q'(s) = \xi'(s+q) - \xi'(q)$ and $\xi_q''(s) = \xi''(s+q)$, we get

$$z_- \stackrel{d}{=} -2(\xi'(1) - \xi'(q)) + \int_q^1 \xi''(s)^{1/2}\, dW_s \stackrel{d}{=} -2\sigma^2(q) + \sigma(q)g,$$
$$z_+ \stackrel{d}{=} 2(\xi'(1) - \xi'(q)) + \int_q^1 \xi''(s)^{1/2}\, dW_s \stackrel{d}{=} 2\sigma^2(q) + \sigma(q)g,$$

where $\sigma(q) = (\xi'(1) - \xi'(q))^{1/2}$ and $g \sim N(0,1)$. Since $\partial_x \Phi_{a,\zeta}(0, \Psi(a,\zeta)) = 0$, the equation (13.13) with $x = \Psi(a,\zeta)$ implies that

$$\mathbb{E}\tanh\big(\Psi(a,\zeta) + z_-\big) \le a \le \mathbb{E}\tanh\big(\Psi(a,\zeta) + z_+\big). \tag{13.14}$$

Note that $\tanh(x)$ is nondecreasing on $\mathbb{R}$ and is concave on $[0,\infty)$, so

$$a \le \mathbb{E}\tanh\big(\Psi(a,\zeta) + z_+\big) \le \mathbb{E}\tanh\big(\Psi(a,\zeta) + |z_+|\big) \le \tanh\big(\Psi(a,\zeta) + \mathbb{E}|z_+|\big),$$

which clearly gives the desired lower bound in (13.12).

The upper bound of (13.12) requires a bit more work. From the left-hand side of (13.14), for any $M > 0$,

$$\begin{aligned}
&\tanh(\Psi(a,\zeta) - 2\sigma^2(0) - \sigma(0)M)\mathbb{P}(|g| \le M) - \mathbb{P}(|g| > M) \\
&\quad \le \mathbb{E}\tanh(\Psi(a,\zeta) - 2\sigma^2(0) - \sigma(0)|g|)I(|g| \le M) \\
&\qquad + \mathbb{E}\tanh(\Psi(a,\zeta) - 2\sigma^2(0) - \sigma(0)|g|)I(|g| > M) \\
&\quad = \mathbb{E}\tanh(\Psi(a,\zeta) - 2\sigma^2(0) - \sigma(0)|g|) \le a.
\end{aligned}$$

Hence,

$$\tanh(\Psi(a,\zeta) - 2\sigma^2(0) - \sigma(0)M) \le \frac{a + \mathbb{P}(|g| > M)}{\mathbb{P}(|g| \le M)}. \tag{13.15}$$

From now on, we choose

$$M = \max\big(8/\sqrt{2\pi}, \sqrt{-2\log(1-a)}\big).$$

First, note that from L'Hôpital's rule,

$$\lim_{a\uparrow 1} \frac{\tanh^{-1} a}{\sqrt{-2\log(1-a)}} = \lim_{a\uparrow 1} \frac{\frac{1}{1-a^2}}{\frac{1}{(1-a)\sqrt{-2\log(1-a)}}} = \lim_{a\uparrow 1} \frac{\sqrt{-2\log(1-a)}}{1+a} = \infty.$$

This means that there exists a constant $c > 0$ such that

$$c\tanh^{-1}(a) \ge \sqrt{-2\log(1-a)}, \quad \forall a \in [0,1). \tag{13.16}$$

Second, this choice of M also implies that

$$\mathbb{P}(|g| > M) \le \frac{2}{M\sqrt{2\pi}} e^{-M^2/2} \le \frac{1-a}{4},$$

where the first inequality is the usual tail bound for the Gaussian random variable. From this,

$$\frac{a + \mathbb{P}(|g| > M)}{\mathbb{P}(|g| \le M)} \le \frac{1+3a}{3+a} < 1$$

so that from (13.15) and (13.16),

$$\begin{aligned} \Psi(a,\zeta) &\le \tanh^{-1}\left(\frac{a + \mathbb{P}(|g| > M)}{\mathbb{P}(|g| \le M)}\right) + 2\sigma^2(0) + \sigma(0)M \\ &\le \tanh^{-1}\left(\frac{1+3a}{3+a}\right) + 2\sigma^2(0) + \sigma(0)M \qquad (13.17) \\ &\le \tanh^{-1}\left(\frac{1+3a}{3+a}\right) + 2\sigma^2(0) + \sigma(0)\left(c\tanh^{-1}(a) + \frac{8}{\sqrt{2}}\right). \end{aligned}$$

To finish our proof, it remains to control the first term. Note that for any $0 \le x \le y < 1$,

$$\begin{aligned} \tanh^{-1}(y) - \tanh^{-1}(x) &= \int_x^y \frac{1}{1-w^2}dw \\ &\le \int_x^y \frac{1}{1-w}dw = -\log\frac{1-y}{1-x}. \end{aligned}$$

In particular, if we take $x = a$ and $y = (1+3a)/(3+a)$, then $0 \le x \le y < 1$ and

$$\begin{aligned} \tanh^{-1}\left(\frac{1+3a}{3+a}\right) - \tanh^{-1}(a) &\le -\log\frac{1-y}{1-x} \\ &= -\log\frac{\frac{3+a-1-3a}{3+a}}{1-a} = -\log\frac{2}{3+a} \le \log 2. \end{aligned}$$

This and (13.17) together complete our proof. □

13.3 Convexity and directional derivative of $\overline{\mathcal{P}}_\mu^v(\lambda,\zeta)$

Finally, we establish two key properties of the functional $(v,\lambda,\zeta) \to \overline{\mathcal{P}}_\mu^v(\lambda,\zeta)$. Let $\mu \in \mathcal{M}_{0,1}$ be fixed and let $q = \int a^2\mu(da)$. For any $\zeta_0,\zeta_1 \in \mathcal{M}_{0,1-q}$, $v_0, v_1 \in \overline{V}$, and $\lambda_0,\lambda_1 \in \mathbb{R}$, denote, for $b \in [0,1]$,

$$\zeta_b = (1-b)\zeta_0 + b\zeta_1, \quad v_b = (1-b)v_0 + bv_1, \quad \lambda_b = (1-b)\lambda_0 + b\lambda_1.$$

First, we show that $(v,\lambda,\zeta) \to \overline{\mathcal{P}}_\mu^v(\lambda,\zeta)$ is convex.

LEMMA 13.4. *We have that*

$$\overline{\mathcal{P}}_\mu^{v_b}(\lambda_b,\zeta_b) \le (1-b)\overline{\mathcal{P}}_\mu^{v_0}(\lambda_0,\zeta_0) + b\overline{\mathcal{P}}_\mu^{v_1}(\lambda_1,\zeta_1).$$

Moreover, if $b \in (0,1)$*, this inequality is strict if either* $\zeta_0 \neq \zeta_1$ *or* $\lambda_0 a + v_0(a) \not\equiv \lambda_1 a + v_1(a)$ *on* $\operatorname{supp}(\mu) \cap (0,1)$.

PROOF. This follows immediately from (13.6) and (13.7). □

Remark 13.5. From this lemma and the Lipschitz property (13.4), if the support of μ contains a point in $(0,1)$ then, for any fixed $v \in \overline{V}$, the functional $(\lambda, \zeta) \to \overline{\mathcal{P}}_\mu^v(\lambda, \zeta)$ is a strictly convex and continuous on $\mathbb{R} \times \mathcal{M}_{0,1-q}$. This guarantees the existence of the unique minimizer $(\lambda_\mu^v, \zeta_\mu^v)$ in $\overline{\mathcal{P}}_\mu^v = \inf_{(\lambda,\zeta)\in\mathbb{R}\times\mathcal{M}_{0,1-q}} \overline{\mathcal{P}}_\mu^v(\lambda, \zeta)$.

Next, we show that the directional derivative of $(v, \zeta) \to \overline{\mathcal{P}}_\mu^v(0, \zeta)$ exists and write down an explicit expression under a certain assumption. Recall the optimal process $u_{a,\xi,x}$ from (13.3).

LEMMA 13.6. *Let* $\zeta_0, \zeta_1 \in \mathcal{M}_{0,1-q}$ *and* $v_0, v_1 \in \overline{V}$. *Assume that* $v_0 = v_1$ *on* $[\delta, 1)$ *for some* $\delta \in (0,1)$. *Then we have that*

$$\begin{aligned}
&\frac{d}{db}\overline{\mathcal{P}}_\mu^{v_b}(0,\zeta_b)\Big|_{b=0+} \\
&= \frac{1}{2}\int_0^{1-q} \xi_q''(s)(\zeta_1(s)-\zeta_0(s))\Big(\int_{[0,1)} \mathbb{E}u_{a,\zeta_0,v_0(a)}(s)^2\mu(da) - s\Big)ds \\
&\quad + \int_{[0,1)} (v_1(a)-v_0(a))\partial_x\Phi_{a,\zeta_b}(0,v_0(a))\mu(da).
\end{aligned} \tag{13.18}$$

PROOF. Recall from (13.8) that, for any $a \in [0,1)$ and $b \in [0,1)$,

$$\begin{aligned}
\frac{d}{db}\Phi_{a,\zeta_b}(0,v_b(a)) &= \frac{1}{2}\int_0^{1-q} \xi_q''(s)\big(\zeta_1(s)-\zeta_0(s)\big)\mathbb{E}u_{a,\zeta_b,v_b(a)}(s)^2\,ds \\
&\quad + \big(v_1(a)-v_0(a)\big)\partial_x\Phi_{a,\zeta_0}(0,v_b(a)),
\end{aligned}$$

where $u_{a,\zeta_b,v_b(a)}$ is the process in (13.2) with $x = v_b(a)$ and $\zeta = \zeta_b$. Furthermore, the right derivative of $\Phi_{a,\zeta_b}(0,v_b(a))$ also exists at $b = 0$ and is equal to the right hand side of the above equation with $b = 0$. From the assumption $v_0 = v_1$ on $[\delta,1)$, we have that from (13.1) and (13.12),

$$\sup_{a\in[0,1),b\in[0,1]} \big|\big(v_1(a)-v_0(a)\big)\partial_x\Phi_{a,\zeta_b}\big(0,v_b(a)\big)\big| < \infty.$$

From this uniform upper bound and the bounded convergence theorem, it follows that

$$\begin{aligned}
&\frac{d}{db}\overline{\mathcal{P}}_\mu^{v_b}(0,\zeta_b)\Big|_{b=0+} \\
&= \lim_{b\downarrow 0}\int_{[0,1)} \frac{\Phi_{a,\zeta_b}(0,v_b(a)) - \Phi_{a,\zeta_0}(0,v_0(a))}{b}\mu(da) \\
&\quad - \frac{1}{2}\int_0^{1-q} \xi_q''(s)s(\zeta_1(s)-\zeta_0(s))ds =
\end{aligned}$$

$$= \frac{1}{2}\int_{[0,1)}\left(\int_0^{1-q}\xi_q''(s)(\zeta_1(s)-\zeta_0(s))\mathbb{E}u_{a,\zeta_0,v_0(a)}(s)^2\,ds\right)\mu(da)$$
$$+\int_{[0,1)}(v_1(a)-v_0(a))\partial_x\Phi_{a,\zeta_0}(0,v_0(a))\mu(da)$$
$$-\frac{1}{2}\int_0^{1-q}\xi_q''(s)s(\zeta_1(s)-\zeta_0(s))ds.$$

Finally, combining the first and third equations together by using Fubini's theorem, the above uniform upper bound, and (13.12) completes our proof. □

Remark 13.7. Suppose that ζ_0 is a minimizer of the variational formula

$$\inf_{\zeta\in\mathcal{M}_{0,1-q}}\overline{\mathcal{P}}_\mu^{v_\zeta}(0,\cdot).$$

In a standard manner as [16, Proposition 1], [33, Proposition 1.1], or [48, Lemma 4.13], by using Fubini's theorem and the above uniform upper bounds of $\partial_x\Phi_{a,\zeta}$ and $\partial_{xx}\Phi_{a,\zeta}$ in (13.1), the directional derivative (13.6) and the minimality of ζ_0 together yield that we must have

$$\int_{[0,1)}\mathbb{E}\big(\partial_x\Phi_{a,\zeta_0}(s,X_{a,\zeta_0,v_{\zeta_0}(a)}(s))\big)^2\mu(da)=s,$$
$$\xi_q''(s)\int_{[0,1)}\mathbb{E}\big(\partial_{xx}\Phi_{a,\zeta_0}(s,X_{a,\zeta_0,v_{\zeta_0}(a)}(s))\big)^2\mu(da)\le 1,$$

for any s in the support of ζ_0.

Acknowledgment. Wei-Kuo Chen was partially supported by NSF grant DMS-16-42207 and grant DMS-17-52184, Dmitry Panchenko was partially supported by NSERC, and Eliran Subag was supported by the Simons Foundation.

Bibliography

[1] Aizenman, M.; Sims, R.; Starr, S. L. Mean-field spin glass models from the cavity-ROSt perspective. *Prospects in mathematical physics*, 1–30. Contemp. Math., 437. Amer. Math. Soc., Providence, RI, 2007. doi:10.1090/conm/437/08422

[2] Auffinger, A.; Ben Arous, G.; Černý, J. Random matrices and complexity of spin glasses. *Comm. Pure Appl. Math.* **66** (2013), no. 2, 165–201. doi:10.1002/cpa.21422

[3] Auffinger, A.; Chen, W.-K. On properties of Parisi measures. *Probab. Theory Related Fields* **161** (2015), no. 3-4, 817–850. doi:10.1007/s00440-014-0563-y

[4] Auffinger, A.; Chen, W.-K. The Parisi formula has a unique minimizer. *Comm. Math. Phys.* **335** (2015), no. 3, 1429–1444. doi:10.1007/s00220-014-2254-z

[5] Auffinger, A.; Chen, W.-K. Parisi formula for the ground state energy in the mixed p-spin model. *Ann. Probab.* **45** (2017), no. 6B, 4617–4631. doi:10.1214/16-AOP1173

[6] Auffinger, A.; Chen, W.-K. On concentration properties of disordered Hamiltonians. *Proc. Amer. Math. Soc.* **146** (2018), no. 4, 1807–1815. doi:10.1090/proc/13864

[7] Auffinger, A.; Jagannath, A. On spin distributions for generic p-spin models. *J. Stat. Phys.* **174** (2019), no. 2, 316—332. doi:10.1007/s10955-018-2188-5

[8] Auffinger, A.; Jagannath, A. Thouless-Anderson-Palmer equations for generic p-spin glasses. *Ann. Probab.* **47** (2019), no. 4, 2230–2256. doi:10.1214/18-AOP1307
[9] Belius, D.; Kistler, N. The TAP-Plefka variational principle for the spherical SK model. *Comm. Math. Phys.* **367** (2019), no. 3, 991–1017. doi:10.1007/s00220-019-03304-y
[10] Ben Arous, G.; Subag, E.; Zeitouni, O. Geometry and temperature chaos in mixed spherical spin glasses at low temperature: the perturbative regime. *Comm. Pure Appl. Math.* **73** (2020), no. 8, 1732–-1828. doi:10.1002/cpa.21875
[11] Bolthausen, E. An iterative construction of solutions of the TAP equations for the Sherrington-Kirkpatrick model. *Comm. Math. Phys.* **325** (2014), no. 1, 333–366. doi:10.1007/s00220-013-1862-3
[12] Chatterjee, S. Spin glasses and Stein's method. *Probab. Theory Related Fields* **148** (2010), no. 3-4, 567–600. doi:10.1007/s00440-009-0240-8
[13] Chatterjee, S. *Superconcentration and related topics.* Springer Monographs in Mathematics. Springer, Cham, 2014. doi:10.1007/978-3-319-03886-5
[14] Chen, W.-K. The Aizenman-Sims-Starr scheme and Parisi formula for mixed p-spin spherical models. *Electron. J. Probab.* **18** (2013), no. 94, 14 pp. doi:10.1214/EJP.v18-2580
[15] Chen, W.-K. Chaos in the mixed even-spin models. *Comm. Math. Phys.* **328** (2014), no. 3, 867–901. doi:10.1007/s00220-014-1888-1
[16] Chen, W.-K. Variational representations for the Parisi functional and the two-dimensional Guerra-Talagrand bound. *Ann. Probab.* **45** (2017), no. 6A, 3929–3966. doi:10.1214/16-AOP1154
[17] Chen, W.-K.; Gamarnik, D.; Panchenko, D.; Rahman, M. Suboptimality of local algorithms for a class of max-cut problems. *Ann. Probab.* **47** (2019), no. 3, 1587–1618. doi:10.1214/18-AOP1291
[18] Chen, W.-K.; Handschy, M.; Lerman, G. On the energy landscape of the mixed even p-spin model. *Probab. Theory Relat. Fields* **171** (2018), no. 1-2, 53–95. doi:10.1007/s00440-017-0773-1
[19] Chen, W.-K.; Panchenko, D. On the TAP free energy in the mixed p-spin models. *Comm. Math. Phys.* **362** (2018), no. 1, 219–252. doi:10.1007/s00220-018-3143-7
[20] Chen, W.-K.; Panchenko, D.; Subag, E. The generalized TAP free energy II. *Comm. Math. Phys.* **381** (2021), no. 1, 257-–291. doi:10.1007/s00220-020-03887-x
[21] Crisanti, A.; Leuzzi, L.; Rizzo, T. The complexity of the spherical p-spin spin glass model, revisited. *The European Physical Journal B - Condensed Matter and Complex Systems* **36** (2003), no. 1, 129–136. doi:10.1140/epjb/e2003-00325-x
[22] Crisanti, A.; Leuzzi, L.; Rizzo, T. Complexity in mean-field spin-glass models: Ising p-spin. *Phys. Rev. B* **71** (2005), 094202. doi:10.1103/PhysRevB.71.094202
[23] Crisanti, A.; Sommers, H.-J. Thouless-Anderson-Palmer Approach to the Spherical p-Spin Spin Glass Model. *J. Phys. I France* **5** (1995), no. 7, 805–813. doi:10.1051/jp1:1995164
[24] Ding, J.; Eldan, R.; Zhai, A. On multiple peaks and moderate deviations for the supremum of a Gaussian field. *Ann. Probab.* **43** (2015), no. 6, 3468–3493. doi:10.1214/14-AOP963
[25] El Alaoui, A.; Montanari, A.; Sellke, M. Optimization of mean-field spin glasses. *Ann. Probab.* **49** (2021), no. 6, 2922–2960. doi:10.1214/21-aop1519
[26] Gamarnik, D.; Jagannath, A. The overlap gap property and approximate message passing algorithms for p-spin models. *Ann. Probab.* **49** (2021), no. 1, 180-–205. doi:10.1214/20-AOP1448
[27] Gamarnik, D.; Jagannath, A.; Wein, A. S. Low-degree hardness of random optimization problems. *2020 IEEE 61st Annual Symposium on Foundations of Computer Science*, 131-–140. IEEE Computer Soc., Los Alamitos, CA.
[28] Guerra, F. Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.* **233** (2003), no. 1, 1–12. doi:10.1007/s00220-002-0773-5
[29] Guerra, F.; Toninelli, F. The thermodynamic limit in mean field spin glass models. *Commun. Math. Phys.* **230** (2002), no. 1, 71–79. doi:10.1007/s00220-002-0699-y

[30] Jagannath, A. Approximate ultrametricity for random measures and applications to spin glasses. *Comm. Pure Appl. Math.* **70** (2017), no. 4, 611–664. doi:10.1002/cpa.21685
[31] Jagannath, A.; Sen, S. On the unbalanced cut problem and the generalized Sherrington-Kirkpatrick model. *Ann. Inst. Henri Poincaré D* **8**(2021), no. 1, 35-–88. doi:10.4171/aihpd/97
[32] Jagannath, A.; Tobasco, I. A dynamic programming approach to the Parisi functional. *Proc. Amer. Math. Soc.* **144** (2016), no. 7, 3135–3150. doi:10.1090/proc/12968
[33] Jagannath, A.; Tobasco, I. Some properties of the phase diagram for mixed p-spin glasses. *Probab. Theory Related Fields* **167** (2017), no. 3-4, 615–672. doi:10.1007/s00440-015-0691-z
[34] Kabashima, Y.; Krzakala, F.; Mézard, M.; Sakata, A.; Zdeborová, L. Phase transitions and sample complexity in Bayes-optimal matrix factorization. *IEEE Trans. Inform. Theory* **62** (2016), no. 7, 4228–4265. doi:10.1109/TIT.2016.2556702
[35] Mézard, M.; Parisi, G.; Sourlas, N.; Toulouse, G.; Virasoro, M. On the nature of the spin-glass phase. *Phys. Rev. Lett.* **52** (1984), 1156. doi:10.1103/PhysRevLett.52.1156
[36] Mézard, M.; Parisi, G.; Sourlas, N.; Toulouse, G.; Virasoro, M. Replica symmetry breaking and the nature of the spin-glass phase. *J. Physique* **45** (1984), no. 5, 843–854. doi:10.1051/jphys:01984004505084300
[37] Mézard, M.; Parisi, G.; Virasoro, M. A. *Spin glass theory and beyond*. World Scientific Lecture Notes in Physics, 9. World Scientific Publishing, Teaneck, NJ, 1987.
[38] Mézard, M.; Virasoro, M. A. The microstructure of ultrametricity. *J. Physique* **46** (1985), no. 8, 1293–1307. doi:10.1051/jphys:019850046080129300
[39] Montanari, A. Optimization of the Sherrington-Kirkpatrick Hamiltonian. *2019 IEEE 60th Annual Symposium on Foundations of Computer Science*, 1417-–1433. IEEE Comput. Soc. Press, Los Alamitos, CA.
[40] Montanari, A.; Richard, E. Non-negative principal component analysis: message passing algorithms and sharp asymptotics. *IEEE Trans. Inform. Theory* **62** (2016), no. 3, 1458–1484. doi:10.1109/TIT.2015.2457942
[41] Montanari, A.; Venkataramanan, R. Estimation of low-rank matrices via approximate message passing. *Ann. Statist.* **49** (2021), no. 1, 321—345. doi:10.1214/20-AOS1958
[42] Panchenko, D. Free energy in the generalized Sherrington-Kirkpatrick mean field model. *Rev. Math. Phys.* **17** (2005), no. 7, 793–857. doi:10.1142/S0129055X05002455
[43] Panchenko, D. The Parisi ultrametricity conjecture. *Ann. of Math. (2)* **177** (2013), no. 1, 383–393. doi:10.4007/annals.2013.177.1.8
[44] Panchenko, D. *The Sherrington-Kirkpatrick model*. Springer Monographs in Mathematics. Springer, New York, 2013. doi:10.1007/978-1-4614-6289-7
[45] Panchenko, D. The Parisi formula for mixed p-spin models. *Ann. Probab.* **42** (2014), no. 3, 946–958. doi:10.1214/12-AOP800
[46] Panchenko, D. The free energy in a multi-species Sherrington-Kirkpatrick model. *Ann. Probab.* **43** (2015), no. 6, 3494–3513. doi:10.1214/14-AOP967
[47] Panchenko, D. Chaos in temperature in generic $2p$-spin models. *Comm. Math. Phys.* **346** (2016), no. 2, 703–739. doi:10.1007/s00220-016-2585-z
[48] Panchenko, D. *Bonus chapter to "The Sherrington-Kirkpatrick model"*. 2018. Available at: `https://sites.google.com/site/panchenkomath/`
[49] Panchenko, D. Free energy in the mixed p-spin models with vector spins. *Ann. Probab.* **46** (2018), no. 2, 865–896. doi:10.1214/17-AOP1194
[50] Panchenko, D. Free energy in the Potts spin glass. *Ann. Probab.* **46** (2018), no. 2, 829–864. doi:10.1214/17-AOP1193
[51] Parisi, G. Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.* **43** (1979), 1754–1756. doi:10.1103/PhysRevLett.43.1754
[52] Parisi, G. A sequence of approximate solutions to the S-K model for spin glasses. *J. Phys. A* **13** (1980), L–115.

[53] Plefka, T. Convergence condition of the TAP equation fo the infinite-ranged Ising spin glass model. *J. Phys. A: Math. Gen.* **15** (1982), no. 6, 1971–1978. doi:10.1088/0305-4470/15/6/035
[54] Rieger, H. The number of solutions of the Thouless-Anderson-Palmer equations for p-spin-interaction spin glasses. *Phys. Rev. B* **46** (1992), 14655–14661. doi:10.1103/PhysRevB.46.14655
[55] Ruelle, D. A mathematical reformulation of Derrida's REM and GREM. *Comm. Math. Phys.* **108** (1987), no. 2, 225–239.
[56] Sherrington, D.; Kirkpatrick, S. Solvable model of a spin glass. *Phys. Rev. Lett.* **35** (1975), 1792–1796. doi:10.1103/PhysRevLett.35.1792
[57] Subag, E. The complexity of spherical p-spin models—a second moment approach. *Ann. Probab.* **45** (2017), no. 5, 3385–3450. doi:10.1214/16-AOP1139
[58] Subag, E. The geometry of the Gibbs measure of pure spherical spin glasses. *Invent. Math.* **210** (2017), no. 1, 135–209. doi:10.1007/s00222-017-0726-4
[59] Subag, E. Free energy landscapes in spherical spin glasses. Preprint, 2018. arXiv:1804.10576 [math.PR]
[60] Subag, E. Following the ground states of full-RSB spherical spin glasses. *Comm. Pure Appl. Math.* **74** (2021), no. 5, 1021-–1044. doi:10.1002/cpa.21922
[61] Subag, E.; Zeitouni, O. The extremal process of critical points of the pure p-spin spherical spin glass model. *Probab. Theory Related Fields* **168** (2017), no. 3-4, 773–820. doi:10.1007/s00440-016-0724-2
[62] Talagrand, M. The Parisi formula. *Ann. of Math. (2)* **163** (2006), no. 1, 221–263. doi:10.4007/annals.2006.163.221
[63] Talagrand, M. Parisi measures. *J. Funct. Anal.* **231** (2006), no. 2, 269–286. doi:10.1016/j.jfa.2005.03.001
[64] Talagrand, M. Construction of pure states in mean field models for spin glasses. *Probab. Theory Related Fields* **148** (2010), no. 3-4, 601–643. doi:/10.1007/s00440-009-0242-6
[65] Talagrand, M. *Mean field models for spin glasses. Volume I*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, 54. Springer, Berlin, 2011. doi:10.1007/978-3-642-15202-3
[66] Talagrand, M. *Mean field models for spin glasses. Volume II. Advanced replica-symmetry and low temperature.*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, 55. Springer, Heidelberg, 2011.
[67] Thouless, D. J.; Anderson, P. W.; Palmer, R. G. Solution of 'Solvable model of a spin glass'. *Philosophical Magazine* **35** (1977), no. 3, 593–601. doi:10.1080/14786437708235992
[68] Zdeborová, L.; Krzakala, F. Statistical physics of inference: thresholds and algorithms. *Adv. Phys.* **65** (2016), no. 5, 453–552. doi:10.1080/00018732.2016.1211393

WEI-KUO CHEN
School of Mathematics
University of Minnesota
537 Vincent Hall, 206 Church St. SE
Minneapolis, MN 55455
USA
E-mail: wkchen@umn.edu

ELIRAN SUBAG
Department of Mathematics
Weizmann Institute of Science
POB 26, Rehovot 76100
ISRAEL
E-mail: eliran.subag@
weizmann.ac.il

Received September 2020.

DMITRY PANCHENKO
Department of Mathematics
University of Toronto
40 St. George St.
Toronto, ON, M5S 2E4
CANADA
E-mail: panchenk@
math.toronto.edu